OXFORD MEDICAL PUBLICATIONS

APPLIED PHYSIOLOGY

BY

SAMSON WRIGHT
M.D., F.R.C.P.

JOHN ASTOR PROFESSOR OF PHYSIOLOGY, UNIVERSITY OF LONDON, MIDDLESEX HOSPITAL MEDICAL SCHOOL. SOMETIME EXAMINER IN PHYSIOLOGY TO THE UNIVERSITIES OF OXFORD, LONDON AND LEEDS; THE ROYAL COLLEGE OF SURGEONS OF ENGLAND; THE ROYAL COLLEGE OF SURGEONS OF EDINBURGH; THE CONJOINT BOARD IN ENGLAND; THE CONJOINT BOARD IN IRELAND

WITH THE COLLABORATION OF

MONTAGUE MAIZELS, M.D., F.R.C.P.

PROFESSOR OF CLINICAL PATHOLOGY, UNIVERSITY OF LONDON, UNIVERSITY COLLEGE HOSPITAL MEDICAL SCHOOL

AND

JOHN B. JEPSON, M.A., B.Sc., D.Phil., A.R.I.C.

SENIOR LECTURER IN BIOCHEMISTRY, COURTAULD INSTITUTE OF BIOCHEMISTRY, MIDDLESEX HOSPITAL MEDICAL SCHOOL

NINTH EDITION

GEOFFREY CUMBERLEGE
OXFORD UNIVERSITY PRESS
LONDON NEW YORK TORONTO

Oxford University Press, Amen House, London E.C.4

GLASGOW NEW YORK TORONTO MELBOURNE WELLINGTON
BOMBAY CALCUTTA MADRAS KARACHI CAPE TOWN IBADAN

Geoffrey Cumberlege, Publisher to the University

First Edition	*1926*
Second Edition	*1928*
Third Edition	*1929*
Fourth Edition	*1931*
Second Impression	*1932*
Fifth Edition	*1934*
Second Impression	*1934*
Third Impression	*1935*
Sixth Edition	*1936*
Second Impression	*1937*
Third Impression	*1938*
Seventh Edition	*1940*
Second Impression	*1941*
Third Impression	*1942*
Fourth Impression	*1943*
Eighth Edition	*1945*
Second Impression	*1947*
Third Impression	*1948*
Ninth Edition	*1952*
Second Impression	*1953*
Third Impression	*1955*

PRINTED IN GREAT BRITAIN BY
MORRISON AND GIBB LIMITED, LONDON AND EDINBURGH

M. W.

IN LOVING MEMORY

Rabbi Akiba (in Roman captivity) to his favourite pupil Simeon ben Yochai: "My son, more than the calf wishes to suck does the cow yearn to suckle."

PREFACE TO THE NINTH EDITION

This edition of *Applied Physiology* is virtually a new book; more than half the text has been rewritten and the rest has been carefully revised to reflect the present state of knowledge. The subject matter has been radically rearranged. The opening chapter now deals with the "Regulation of the Internal Environment," a theme made classical by Claude Bernard and developed by many great physiologists like Sherrington, Haldane, and Barcroft in this country, and Cannon, Lawrence Henderson, and Van Slyke in the United States. The "constancy of the internal environment" is the outstanding generalization of physiology; the first 130 pages of this book, which illustrate the principle by many examples, have enabled me to review under one heading numerous themes which are usually treated apart because they fall into what conventional classification treats as separate systems. On the clinical side this chapter deals with many topics of practical importance, such as water and salt lack and excess; tests of renal function; renal failure; acidæmia and alkalæmia; diabetes insipidus; œdema; blood changes in obstruction of the alimentary canal; disorders of the cerebrospinal fluid. Other newly arranged chapters are those on "Metabolism," "Endocrine Control of Metabolism" and "Reproduction."

An important feature of this edition is an increase in the number of illustrations. About 200 new figures have been added, bringing the total up to 688; these figures have been carefully chosen, mainly from recent papers; most of them have been redrawn or modified to make them more suitable for teaching purposes. My experience makes me doubt whether illustrations are adequately or properly used. Many welcome them as mere padding material which leaves less text on the page and enables one consequently to cover the ground more quickly. It is curious that though diagrams are studied with care, illustrations which summarize the results of actual experiments and so represent first-hand evidence are merely glanced at. I again draw attention to the Appendix (p. 1118) in which two figures are carefully described and analysed. It might be worth the reader's while to examine many of the illustrations in the text in the same thorough way. The sources of all the figures are fully acknowledged in the text; I might add that any paper that contains a figure worth borrowing is invariably well worth reading. Professor Diamond and the Commonwealth Fund of New York kindly gave me permission to prepare the three new colour plates from the beautiful illustrations in the book by Blackfan, Diamond, and Leister, *Atlas of the Blood in Children.*

As in the past, the references have been restricted to a few key papers which I have found especially useful; they will serve to introduce the reader to the copious literature of the subject.

The terminology in the chapters on the "Nervous System" and "Autonomic Nervous System" with few exceptions follows that used by

Ranson and Clark in their *Anatomy of the Nervous System*; but mention is made in brackets of the older terms which still have a wide currency.

When this book first appeared more than twenty-five years ago it was something of a pioneer venture. The point of view which it advocated was set out in the original Preface (which has been regularly reprinted) and illustrated by the text itself. In selecting the material for discussion I have always had in mind the needs and interests of both undergraduate students and postgraduate readers, including those preparing for the higher examinations, who wish to keep abreast with current developments in both pure and applied physiology. The subject matter is presented in considerable detail because the details are needed for a proper understanding of the workings of the body in health and disease and are often of great practical importance in the care of patients; but I leave it to the good sense of the individual reader to concentrate on his needs and to ignore what he considers to be outside his field of interest. In this way he can make the book as long or as short as he wishes.

More than ever before I am indebted to friends and colleagues for their help, especially to my collaborators Professor Montague Maizels and Dr. John B. Jepson. Professor Maizels has advised and assisted me in the preparation of the chapters on the "Internal Environment" and the "Blood," and on matters relating to Clinical Pathology. Dr. Jepson has rewritten the chapter on "Metabolism" and has advised and assisted me on matters relating to biochemistry and physical chemistry throughout the book; his contribution has been specially notable in the chapters dealing with the "Endocrines" and "Nutrition." Dr. Cyril Keele has helped with the "Endocrines" and "Reproduction." Mr. W. F. Floyd has contributed the section on "Clinical Electroencephalography" and has advised on electrophysiological matters. I gratefully acknowledge help from: Dr. Moran Campbell, Miss Mary Chennells, Dr. C. E. Dent, Dr. P. W. Nathan, Dr. Eric Neil, and Professor David Slome. Special thanks are due to Dr. I. Calma, who has been good enough to prepare the Index.

The publishers, and especially Mr. G. T. Hollis, M.A., have been unfailing in their kindness and consideration. They have agreed readily to all my suggestions for improving the book. They have permitted without impatience or hint of complaint most extensive and repeated alterations in the proofs; by their indulgence I have been able to incorporate in the text references to the major advances in knowledge which took place while the book was passing through the press and to make the consequential alterations needed to maintain self-consistency. My daughter Sandra Wright devoted much skill and patience to transmuting a long and difficult manuscript into a clear and intelligible typescript. My secretary Miss R. Bambridge has given me devoted service in the long and complex task of seeing the book through the press. My warmest thanks are extended to them all.

S. W.

MIDDLESEX HOSPITAL MEDICAL SCHOOL,
January, 1952.

FROM THE PREFACE TO THE FIRST EDITION

This book is based on the various courses of lectures I have given both in the department of Physiology in the Middlesex Hospital Medical School and to men working in the wards and the various departments of the Hospital. My aim has been to present clearly and concisely how the various functions of the body are performed. To economize space, no account is given, however, of the peripheral mechanism of vision or of hearing; otherwise, the ground is covered fairly completely. As the book is intended for medical readers, I have throughout laid emphasis on and devoted particular attention to those functions which are commonly deranged in disease. I have kept in view Dr. Haldane's dictum, that the aim of physiology is to consider how the internal environment of the body is kept constant in spite of continual alterations in the external environment. Throughout the text numerous cross references have been inserted which may help to clear up doubtful points, and are especially intended to impress on the reader that the subject must be considered as a coordinated and integrated whole, and that it does not deal with a series of isolated and independent organs and systems.

I have not attempted, in the more strictly applied sections of the book, to cover the ground completely, but I have selected from the mass of material available the problems of the greatest practical importance, and particularly those which are, perhaps, not so fully discussed in text-books of pathology or medicine.

I must confess to have paid some attention to the requirements of the various examining bodies. The physiological sections may prove useful to men reading for the second M.B. examination of the various Universities, and in particular I hope it will be helpful to men preparing for the Primary Fellowship examination of the Royal College of Surgeons. The applied sections may be found helpful by men reading for their final examinations, and perhaps even for the higher examinations of M.D. or M.R.C.P.

It seems to me unfortunate that in the teaching of physiology greater use is not made of the wealth of clinical material present in the wards of the Hospitals, which could readily be made available. The main facts of the physiology of the nervous system and the ductless glands, at any rate, could thus be clearly demonstrated; and the interest of the student would be aroused when he finds that he is considering in his physiological studies the same patients that he will have to deal with when doing his clinical work. This is the method I have adopted whenever possible in this book. In considering the cerebellum, for example, a full account is given of the symptoms and signs of lesions to this structure in the *human* subject. This enables the student of physiology to draw his conclusions concerning the functions of the organ, and at the same time familiarizes him with the clinical

aspects of disorders of the cerebellum. If a course of demonstrations such as I have suggested became an integral part of all courses in physiology, the medical student would enter on his clinical studies with a much sounder knowledge of the practically important parts of the subject, and with at least an elementary knowledge of medicine.

S. W.

MIDDLESEX HOSPITAL MEDICAL SCHOOL
September, 1926.

CONTENTS

NOTE

As was explained in the Preface, some passages in this book are wholly or mainly of *clinical or specialist interest.* In order to guide the pre-clinical student in his reading, a list of such passages, with page references, is given below. These may either be wholly *omitted* by him, or studied with discretion. Particulars are also given of topics which are treated in considerably greater detail than he needs; in these cases, he might be wise to try to grasp only the main outlines of the discussion.

PAGE	
7–9.	Cation changes in stored blood.
70–77.	Kidney function in disease.
78–79.	Artificial kidney.
86–87.	Blood volume in disease.
99–103.	Pathological changes in blood reaction.
103–110.	Effects on body fluids of obstruction of alimentary canal.
110–118.	Œdema.
125–132.	Pathology of cerebrospinal fluid.
147–150.	Intravascular thrombosis.
157–160.	Purpura.
169–170.	Fragility of blood cells in disease.
170–171.	Action of hæmolysins.
172–173.	Sedimentation rate.
184–186.	Use of stored blood.
190–193.	Jaundice.
213–214.	Hæmochromatosis.
251–255.	ECG changes in myocardial lesions.
255–256.	Axis deviation.
267.	Cardiac murmurs.
282–295.	Abnormalities of cardiac rhythm.
295–300.	Heart failure. Circulation in heart disease.
333–335.	Patent ductus arteriosus.
339–344.	Traumatic shock.
345–358.	Experimental and clinical hypertension.
358–361.	Disorders of peripheral circulation.
368–372.	Pneumothorax.
446–450.	Types of anoxia.
451–455.	Cyanosis.
455–465.	Dyspnœa.
564–566.	Detailed anatomy of thalamus.
567.	Thalamic syndrome.
579.	Argyll Robertson pupil.
612.	Clinical cerebellar dysfunction.
624–629.	Clinical electroencephalography.
629–646.	The account of excitomotor areas, pyramidal tracts, and clinical hemiplegia is too full.
652–653.	Disturbances of voluntary movement.
653–656.	Speech.
660–661.	Clinical affections of corpus striatum. Chorea and athetosis.
661–668.	Most of the details of the physiology of emotions.
673–675.	Prefrontal leucotomy.
675–685.	Most of the details of conditioned reflexes.
704.	Local lesions of brain stem.
747–765.	A good deal of this passage, dealing with visceral sensibility in disease.
771–773.	Disorders of micturition.
803.	Removal of gall-bladder.
804–805.	Biliary obstruction.
815–818.	Disorders of movements of the alimentary canal.
822–823.	Portal obstruction.
829.	Clinical hepatic failure.
829–833.	Clinical tests of liver function.
833.	Acute hepatic failure.
869.	Fatty liver.
898–899.	Gout.
916.	Hyperinsulinism.
925–927.	Administration of, and abnormal responses to insulin.
940–941.	Clinical pituitary syndromes.
951–954.	Clinical indices of corticoid secretion.
960–964.	The account of the actions of adrenal corticoids (including DOCA and cortisone) is too full.
965–969.	The account of Cushing's disease and adrenogenital syndrome is too full.
989–997.	The account of Graves' disease and its treatment is too full.
1012.	Osteomalacia.
1017–1018.	Hyperplasia of thymus and status thymo-lymphaticus.
1031–1036.	Singapore captivity.
1049–1050.	Clinical undernutrition.
1060–1064.	British rationing system.
1082.	Therapeutic uses of sex hormones.
1115–1117.	Carcinoma of the prostate.

APPLIED PHYSIOLOGY

I

REGULATION OF CONSTANCY OF INTERNAL ENVIRONMENT [MILIEU INTÉRIEUR]

THE INTERNAL ENVIRONMENT (MILIEU INTÉRIEUR)[1]

Internal Environment (Milieu Intérieur.)—Claude Bernard pointed out that the external environment of the organism as a whole, *e.g.* atmospheric air in our own case, is not the external environment of the individual cells of the body. The cells are carefully insulated from the general external environment and are bathed by the fluid in the minute spaces between the cells—the so-called *tissue* or *interstitial fluid*. Interstitial fluid can normally be collected only with great difficulty; fortunately, however, its crystalloid composition is almost identical with that of plasma which is easily obtained for analysis. It is customary in fact to use plasma analyses as a guide to variations in the composition of the interstitial fluid. Claude Bernard called the interstitial fluid and the plasma the *internal environment* (*milieu intérieur*).[2] He emphasized that in spite of the very wide range of variation that occurs in the composition of the external environment, the internal environment by contrast is kept remarkably constant as a result of the intervention of many compensatory mechanisms. The way in which the constancy of the internal environment is maintained with respect (for example) to acid-base equilibrium, osmotic pressure, concentration of individual solutes or of ions (sugar, Na^+, K^+, Ca^{++}), and temperature is one of the most important problems of physiology. The constancy of the internal environment is not absolute; under normal conditions slight variations occur as with all so-called physiological "constants." If the stresses imposed on the system become too great the composition of the internal environment may alter significantly, often with disastrous effects. Claude Bernard pointed out that the constancy of the composition of the *milieu intérieur* is "the condition of a free life." If by a "free life" is meant vigorous and effective activity of the organism as a whole, the aphorism is extremely apt. Thus a deviation of H^+ ion concentration from the normal pH of 7·4 to 7·0 or 7·8, a change of serum calcium from 10 mg. to 5 or 15 mg–%, a change of blood sugar from 100 to 30 mg–% may cause loss of consciousness and the danger of imminent death.

Living cells are always undergoing change, but to survive and function

[1] Claude Bernard, *Introduction to the Study of Experimental Medicine* (English translation), N.Y., 1949. Olmsted, *Claude Bernard, physiologist*, N.Y., 1938. Barcroft, *Architecture of Physiological Function*, Cambridge, 1934. Cannon, *Wisdom of Body*, N.Y., 1932. Sherrington, *Man on his Nature*, Cambridge, 1951. Sunderman and Boerner, *Normal Values in Clinical Medicine*, Philadelphia, 1949.

[2] "Qu'est-ce que ce milieu intérieur ? C'est le sang, non pas à la vérité le sang tout entier, mais la partie fluide du sang, le plasma sanguin, ensemble de tous liquides interstitiels, source et confluent de tous les échanges élémentaires." (Claude Bernard.)

normally their integrity in all respects, *i.e.* their form, chemical composition and chemical organization must be constantly and fully restored and preserved. Cells maintain the composition of their organic constituents not by inactivity but by ceaseless activity. Innumerable enzymes are constantly breaking down all the organic cell constituents, both large and small; on the other hand synthetic enzymes are constantly making good this "chemical erosion" by rebuilding the cell structure from the components of the diet and from cellular débris. The stability of the living cell is a *dynamic* one, resulting from the establishment of a balance between the breakdown and repair reactions. The most comprehensive generalization of physiology is that the organism reacts to environmental changes (internal or external) in such a manner as to preserve the integrity of the whole organism and of its constituent parts.

The fluid inside the tissue cells is called the *intracellular fluid*; when its composition is compared with that of the surrounding medium (the *interstitial* or *extracellular fluid*) marked differences are revealed (p. 5). To maintain the characteristic pattern of its electrolyte content the cell must use energy which is derived from catabolic processes (p. 8). In order that a cell may function normally it must maintain the constancy of its own internal "private" composition and also be bathed in a surrounding fluid of utterly different composition which also must be kept approximately constant. Thus, if the Na^+ ion concentration in the interstitial fluid is decreased, the excitability of nerve fibre and skeletal and heart muscle diminishes; in the absence of interstitial Na^+ ions, excitability disappears.

The body then is a highly sensitive, highly efficient self-regulating and thus self-preserving mechanism; it automatically preserves the constancy of its own private internal world in the face of the widest variations in the character of the outside world. But with the development of *intelligence*, man need no longer depend exclusively on his inborn unconscious reactions; he can devise new supplementary protective methods. Thus he helps to maintain the constancy of body temperature at the North Pole and Equator by attending to such matters as clothes and shelter; in "conditioned" houses he can even provide for himself the external environment most congenial and convenient to his body and mind. At high altitudes he can overcome the effects of oxygen lack (*anoxia*) not only by automatic physiological reactions but also by deliberately providing himself with a supply of pure oxygen or by encasing himself in a pressurized chamber in which the atmospheric conditions present at sea level are maintained.

BODY WATER AND BODY FLUID. GENERAL SURVEY OF WATER AND FLUID EXCHANGES[1]

Body Water.—Water is the largest constituent of the body; about 65–70% of the total body weight consists of water. The body of a man weighing 70 kg. thus contains about 50 L of water. The water content of most tissues is 70–80%, the skeleton being the principal exception with a water content of about 20%. It is a little surprising that bones should

[1] Gamble, *Extracellular Fluid*, Harvard University Press, 1949. Peters, *Body Water*, Springfield, Illinois, 1935. *Physiol. Rev.*, 1944, *24*, 49.

contain even so much water, but it must be remembered that one-third of the weight of a fresh bone consists of organic material dissolved in water (the inorganic salts constitute the other two-thirds).

Body water is found:

(i) *inside* the cells—*intracellular* water;

(ii) *outside* the cells—*extracellular* water. The extracellular water is further subdivided into: (*a*) water in the *plasma*; (*b*) *interstitial water*, present in the cracks and crevices (*tissue spaces*) which lie between the tissue cells. Additional minor divisions of the extracellular water are: *lymph* (in the lymphatic vessels), *cerebrospinal fluid*, and *aqueous humour* (in the anterior chamber of the eye).

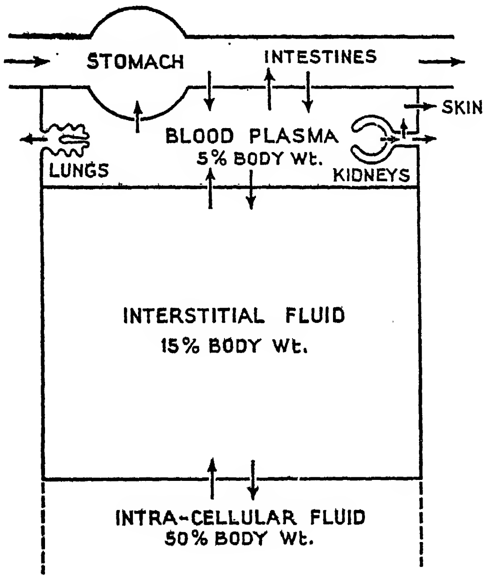


FIG. 1.—Diagram showing Principal Compartments of Body Fluid and Main Fluid Exchanges that take place in the Body. (After Gamble, *Extracellular Fluid,* Harvard University Press, 1949.)

Approximate quantitative data for a man weighing 70 kg. and for one weighing 60 kg. are summarized below (see also Fig. 1):

Total body weight	100%	70 kg.	60 kg.
Total body water	70%	50 L	42 L
(i) Intracellular water	50%	35 L	30 L
(ii) Extracellular water	20%	15 L	12 L
(*a*) plasma	5%	3·5 L	3 L
(*b*) interstitial water	15%	11·5 L	9 L

Water Intake and Water Loss.—Body water is constantly carrying out exchanges with the external environment (Fig. 1).

(1) Water is normally *absorbed* into the body from the *bowel*; it can also be introduced artificially, *e.g.* by subcutaneous or intravenous injection.

(2) Water is *eliminated* from the body:

(i) via the *kidney* in the urine;
(ii) via the *skin* (*a*) in so-called insensible perspiration and (*b*) in sweat;
(iii) via the *lungs* in the expired air;
(iv) to a minor degree via the *large intestine* in the fæces, in lactating women in the *milk* (and from time to time in the *tears*).

(3) It must also be remembered that water is constantly being *formed* in all tissues as an end-product of the oxidation of the foodstuffs.

Approximate quantitative data (both average and range) for these exchanges per day are as follows:

Water intake: as water, 1500 c.c. [range: 0 to several litres per hour];
in food, 1000 c.c. [depending on composition of diet];
from oxidation in tissues, 300 c.c.

Water loss: in urine, 1500 c.c. [range: under 20 c.c. to over 1200 c.c. per hour];
via skin $\begin{cases}\text{insensible perspiration [constant at 600–800 c.c.]}\\ \text{sweat [range: 0 to nearly 2 litres per hour];}\end{cases}$
via lungs, 400 c.c.;
in fæces, 100 c.c. [increased in diarrhœa].

Over a reasonable period of time water intake and water loss must be equal if the normal water balance is to be maintained. Water intake can be reduced to nothing, the only source of water then being the small volume formed during the oxidation of foodstuffs. Water loss, however, cannot normally be reduced to the same degree; thus

(i) insensible perspiration from the skin and water loss by evaporation from the lungs never fall below about 1000 c.c. per day; and
(ii) a minimal flow of urine of about 400 or 500 c.c. per day is needed for the excretion of waste products.

Excessive water loss produces dehydration, and excessive water retention produces hydration; lowered water content of the blood is called *anhydræmia* and excessive water content *hydræmia.*

Body Fluid.—Body water plus its dissolved solutes is called body fluid. Body fluid can be regarded as lying in three *compartments* (Fig. 1):

(i) the *intracellular* fluid which is bounded by the membranes of the individual cells;
(ii) the *plasma* which is bounded by the walls of the blood vessels, and specifically in the capillaries by a thin-walled endothelium;
(iii) the *interstitial* fluid (*tissue* fluid) which is separated from the intracellular water by the cell membranes and from plasma by the capillary endothelium.

Composition of Body Fluids.—Considered initially in a simplified way, *plasma* consists of water, dissolved crystalloids (the chief being NaCl) and dissolved colloids (mainly protein, about 7–8%). *Interstitial (tissue) fluid* has approximately the same crystalloid composition as plasma but its protein content is small.[1] *Intracellular fluid* has an entirely different composition: in many cells it contains little or no Na^+ or Cl'; the principal cation is K^+

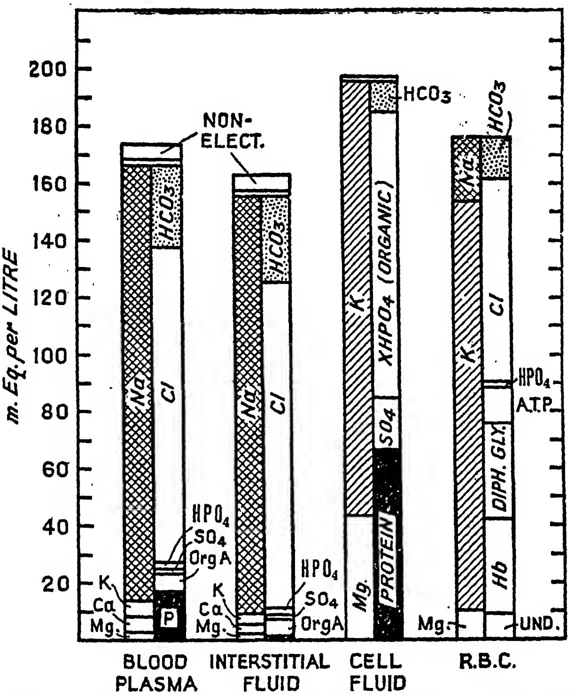


FIG. 2.—*Concentrations* of Constituents of Plasma, Interstitial Fluid, Tissue Cells and Red Blood Corpuscles. (After Gamble, *Extracellular Fluid*, Harvard University Press, 1949.)

Concentrations are expressed in milliequivalents per litre *of water*. Non-Elect. = non-electrolytes; HPO_4 = inorganic phosphate; $XHPO_4$ = organic phosphate; Org. A = organic acids; Und. = undetermined; P = plasma protein; A.T.P. = adenosine triphosphate; Diph. Glyc. = diphosphoglycerate; Hb = hæmoglobin. In each box diagram the left-hand column depicts cations, the right-hand column anions.

(and some Mg^{++}): the principal anions are organic phosphate ($XHPO_4''$), bicarbonate (HCO_3'), and sulphate (SO_4''); it contains far more protein than plasma; in muscle fibres, for example, the protein content is about 20%. The *concentrations* of the constituents of the body fluids are compared in Fig. 2; the values are expressed per litre of *water*. The Table on p. 6 gives the composition of plasma; the values are there expressed per L or per 100 c.c. of *plasma* (93% of the plasma and 70% of the red cells is water) (cf. p. 8).

[1] The composition of cerebrospinal fluid is set out in detail on p. 119.

Composition of Plasma [1]

	Base			Acid	
	m.Eq./L	mg./100 c.c.		m.Eq./L	mg./100 c.c.
Na^+	143	330	HCO_3'	27	165
K^+	5	19·5	Cl'	103	365
Ca^{++}	5	10	HPO_4''	2	3·5
Mg^{++}	2	2·5	SO_4''	1	1·5
	—		Organic Acid	6	
	155		Protein	16	
				—	
				155	

Internal Fluid Exchanges.—In addition to the exchanges between the body and the external environment referred to above there is a much vaster *internal* turnover of water:

(i) Water is secreted in large amounts (about 8–9 L per day) by the digestive glands (salivary glands, stomach, pancreas, liver, small intestine) from the blood into the lumen of the bowel whence it is completely reabsorbed into the blood (almost entirely in the small intestine), apart from the very trivial loss in the fæces.

(ii) A constant large-scale to and fro interchange is always taking place between the plasma (in the vascular compartment) and the interstitial fluid (in the tissue spaces).

(iii) In the kidney some 170 L of water are filtered out in the 24 hours from the glomerular capillaries into the capsules of Bowman to pass down the renal tubules whence it is almost completely reabsorbed into the blood again, apart from about 1500 c.c. which are passed out as urine.

(iv) Smaller scale exchanges are taking place between the blood and the cerebrospinal fluid or the aqueous humour. The principal internal turnovers can be summarized thus:

Bowel: blood→digestive glands→lumen of bowel→blood.
Kidney: blood→glomerular capillaries→Bowman's capsule (glomerular filtrate→down renal tubule→blood.
Capillaries: blood→tissue fluid ↗ blood / ↘ lymphatics→blood.
Brain: blood→choroid plexuses→cerebrospinal fluid→blood.

Interchanges between Various Fluid Compartments.—This question will be dealt with more fully later but must be touched on here.

Plasma and Interstitial Fluid.—The capillary endothelium intervenes between these two fluids. The endothelium is freely permeable to water and all dissolved crystalloids; under conditions of equilibrium therefore the concentration of the crystalloids in plasma and interstitial fluid is roughly the same. Normally, however, the capillary membrane is almost completely impermeable to protein; plasma contains 7–8% of protein while interstitial fluid contains very little. The presence of *ionized* protein molecules in the plasma leads to certain differences, in electrolyte composition, between plasma and interstitial fluid (the so-called *Donnan equilibrium*). In abnormal states

[1] The meaning of m.Eq./L (milliequivalents per litre) is explained on p. 12.

the permeability of the capillary endothelium is increased and then larger concentrations of protein escape into the interstitial fluids.

INTERSTITIAL FLUID AND INTRACELLULAR FLUID.[1]—*Permeability of Cell Membranes.*—The properties of the membranes enclosing tissue cells and red cells have been studied in great detail but our knowledge of the subject is still incomplete. The chief findings are as follows:

(i) The cell membrane is freely permeable to water.

(ii) It is—and has to be if the cell is to live at all—freely permeable to food materials such as glucose, acetoacetate and amino-acids which pass in, and products of metabolism like urea, which pass out. Some metabolites like lactate may move in or out of the cell under different circumstances.

(iii) The cell membrane is presumed to be impermeable to protein; but in the case of the liver, protein can undoubtedly move in and out of the cells to regulate the plasma protein concentration (cf. p. 138).

(iv) The cell membrane is impermeable to *organic* phosphate anions (represented as $XHPO_4''$, where X is the organic radical, *e.g.* in adenosine triphosphate); it is permeable to *inorganic* phosphate ions. The tissue cells (and red cells) are rich in organic phosphate which is thus "locked away" within them; the plasma and interstitial fluid only contain small amounts of inorganic phosphate.

(v) The red cell membrane is freely permeable to other inorganic anions, *e.g.* chloride and bicarbonate (cf. p. 419). Tissue cells (unlike red cells) contain no chloride although the interstitial fluid is rich in chloride (Fig. 2). It is alleged that tissue cell membranes are impermeable to chloride; the true explanation may be that the chloride is actively extruded by the cell as is known to be the case for sodium (*infra*). More information is needed on this subject.

(vi) Tissue cells contain little or no sodium; red cells contain some sodium but in far lower concentration than does plasma. Tissue cells and red cells are rich in potassium, the concentration far exceeding that in plasma. It was thought in the past that cell membranes were impermeable to cations and thus to sodium and potassium. This view is incorrect, as shown by the following observations.

(*a*) Studies with radio-active isotopes show that Na^+ and K^+ move freely in and out of red cells in the circulating blood.

(*b*) During the activity of nerve fibres Na^+ leaks in and K^+ leaks out of the fibre; these movements are reversed during the recovery process (p. 485).

(*c*) Red cells stored in the cold outside the body lose K^+ and gain Na^+; after reinjection into the circulation they expel the newly added Na^+ and regain K^+. It will be shown that living red cells by using energy derived from metabolic processes actively maintain their normal, characteristic cation pattern which is so different from that of the surrounding plasma. When red cell metabolism is depressed, or ceases, the cations move passively with the concentration gradient, *i.e.* the red cells then lose K^+ and gain Na^+.

These phenomena are discussed more fully below.

CATION CHANGES IN STORED BLOOD.—Samples of blood treated as shown in the Table (p. 8) were examined. (i) In fresh blood the red cell cation *contents* are K^+, 100, and Na^+, 12, m.Eq. per litre of *red cells*; since the cells contain about 70% of water, the amount of K^+ in cell *water*, *i.e.* the *concentration* of

[1] Ussing, *Physiol. Rev.*, 1949, *29*, 127 (transport of ions across cellular membranes).

K^+ (usually written $[K^+]$) is 144 m.Eq./L ; $[Na^+]$ is 17 ; $[K^+ + Na^+]$ is 161. In plasma, $[K^+]$ is about 5, and $[Na^+]$ is about 145 m.Eq./L of *plasma*. As plasma contains about 93% water, plasma $[K^+]$ is 5 and $[Na^+]$, 155 m.Eq./L of plasma *water*. Total $[K^+ + Na^+]$ is 160. It will be noted that cells and plasma are in osmotic equilibrium. (Table, (1).)

TABLE (Data from Maizels).

	Volume. μ^3	Cells.				Plasma.			
		Contents m.Eq./L of *cells*.		Concentration m.Eq./L of *cell water*.		Contents m.Eq./L of *plasma*.		Concentration m.Eq./L of *plasma water*.	
		K^+	Na^+	$[K^+]$	$[Na^+]$	K^+	Na^+	$[K^+]$	$[Na^+]$
(1)	85	100	12	144	17	5	145	5	155
(2)	89	70	45	99	63	24	123	26	132
(3)	97	60	62	81	83	30	112	33	123
(4)	86	92	20	131	29	10	139	11	148

Treatment of blood.

1. Freshly drawn blood.
2. Cold-stored for 7 days ; no glucose added.
3. Blood treated as in (2) and then incubated at 37° C. for 18 hours (*no* glucose added).
4. Blood stored as in (2) and then incubated at 37° C. for 18 hours (glucose *added*).

(ii) If fresh blood is cold-stored, red cell metabolism is reduced to a minimum ; as a result the cations flow with the concentration gradients, cell K^+ thus falling and Na^+ rising. As the latter proceeds more rapidly than the former, *total* cell base $(Na^+ + K^+)$ increases ; because of the increased intracellular crystalloid osmotic pressure, water flows into the cell (Table, (2)) and the cell volume rises from 85 to 89 μ^3 in 7 days ; because of this movement of water, the cation *concentrations* in cells and plasma still balance. If the cold-stored blood is now incubated at body temperature *without added glucose*, the blood glucose is soon used up and red cell metabolism ceases. The process of passive diffusion of Na^+ into, and of K^+ out of, the cells with the respective concentration gradients, is speeded up by the rise of temperature. *Total* cell base *content* and cell volume, therefore, increase further although total cell and plasma base (cation) *concentrations* still remain approximately equal (Table, (3)).

(iii) If, on the other hand, *glucose is added* to the cold-stored blood and the mixture is incubated at body temperature, cell metabolism is restored and the cation pattern begins to return to normal ; thus the cell Na^+ concentration is lowered from 63 to 29 m.Eq./L although external, *i.e.* plasma $[Na^+]$ is initially 132 m.Eq./L ; cell $[K^+]$ rises, *e.g.* from 99 to 131 m.Eq./L ; the movements of both Na^+ and K^+ are occurring *against* the concentration gradient. Both total base and cell water decrease and cell volume returns practically to normal (Table, (4)).

Mechanism of Active Cation Transfer.[1]—Movements of ions against concentration gradients require the expenditure of energy, which in the red

[1] For full discussion see Flynn and Maizels, *J. Physiol.*, 1949, *110*, 301 ; Maizels, *J. Physiol.*, 1951, *112*, 59.

cell is supplied by metabolizing glucose; mannose and fructose will also serve, but not galactose or other sugars. The process is not inhibited by cyanide or carbon monoxide and so is independent of the cytochrome mechanism and of respiration; but it is inhibited by sodium fluoride and iodoacetate and so depends on energy supplied by glycolysis, *i.e.* the breakdown of glucose to lactic acid which can take place in the presence, or absence, of oxygen. Thus although human red cells have a negligible level of respiration they can obtain energy for cation transfer from glycolysis.

In the recovering red cell, the energy is used primarily to expel the Na^+ (which has entered), out of the cell into the interstitial fluid. As a result of this process, the physical conditions within the cell are changed in such a manner as to permit the intracellular protein and organic phosphate anions to attract K^+ into the cell from the interstitial fluid against the concentration gradient. In normal red cells, circulating in the plasma, the constant tendency of Na^+ to enter is overcome directly by the energy liberated in the cell, and secondarily the cell K^+ is retained. It is probable that the cation pattern of other tissues is maintained in the same way. Thus the distribution of K^+ and Na^+ in muscle and nerve fibres is similar to that found in human erythrocytes; its normal maintenance likewise depends on active cation transport and probably on glycolysis.

Glycolysis, which is a less effective source of energy than respiration, releases enough energy to enable a cell to maintain its *own* cation pattern. But when a cell is also engaged in mediating active ionic transport on behalf of other tissues, glycolysis does not supply sufficient energy; thus the renal cells which also have to transfer ions from the lumen of the renal tubule into the blood often against the concentration gradient so as to maintain the electrolyte pattern of the whole body, must *respire* to obtain enough energy for this work; if the cells are poisoned with cyanide active ionic transfer ceases (p. 29).

Determination of Volume of Water in the Various Body Compartments.[1]—The principles underlying the methods of determining the volume of plasma, interstitial fluid, and intracellular fluid are simple.

PLASMA VOLUME.—(1) *Dyestuff Method.*[2]—To determine plasma volume a non-toxic dyestuff which is easily identified and readily determined quantitatively is injected intravenously. The ideal substance after intravenous injection should (i) not penetrate into the red blood cells, (ii) not leak out through the capillary endothelium into the tissue spaces, and (iii) be rapidly and uniformly distributed throughout the entire circulation. The degree of dilution of the dye is a measure of the plasma volume. The dye Evans' blue (T1824) has very approximately these properties; it is retained in the circulation because it is bound with serum albumin. The results given by this method tend to err on the high side.

(2) *Radio-iodine-plasma-protein Method.*[3]—5 mg. of iodine containing a trace of radio-active isotope (either $^{130}I^*$, half-life 13 hours or $^{131}I^*$, half-life 8 days) are added to 50 c.c. of plasma *in vitro*; the iodine combines with the plasma protein. The treated plasma is injected intravenously; the degree of dilution of its radio-activity is a measure of the plasma volume. The

[1] See *Methods in Medical Research*, 1951, *4*, Section II, N.Y.
[2] Noble and Gregerson, *J. clin. Investig.*, 1946, *25*, 158.
[3] Fine and Seligman, *J. clin. Investig.*, 1944, *23*, 720.

method is based on the relative impermeability of the capillary endothelium to plasma protein.[1]

RED CELL VOLUME.—(1) *From Plasma Volume and Hæmatocrit Value.*—The relative volume of plasma and corpuscles is determined by centrifuging oxalated blood in a graduated tube (*hæmatocrit*); the corpuscles sink to the bottom and are closely packed together, the plasma rises to the top; the volume of the packed red cells and of the supernatant plasma is read off. The average normal findings are that 100 c.c. of blood contain 45 c.c. of red cells and 55 c.c. of plasma. If the total plasma volume and the ratio of red cell/plasma volume are known, the total red cell volume can be calculated. Thus if the plasma volume is 2750 c.c. and the red cell/plasma volume ratio is 45/55, then the total cell volume is $2750 \times \frac{45}{55} = 2250$ c.c. The calculation assumes that the cell/plasma ratio in the blood in a *large vein* (the blood usually examined) is the same as in the blood volume *as a whole*. This is, however, not the case; when the cell volume is determined directly (*infra*) and is compared with plasma volume, the cell volume is found to be *less* in the ratio of 0·85/1 than that calculated from the plasma volume determination and hæmatocrit findings.[2] This means that the blood in the small vessels contains relatively *fewer* cells than the blood in the main vessels. If the cell/plasma ratio as determined in blood from the main arteries and veins is represented as 1, then the relative cell content elsewhere is: medium vessels 0·92; whole body blood 0·85; minute vessels (arterioles, capillaries, venules) 0·68; liver vessels about 1; spleen vessels over 1.

(2) *Direct Red Cell Volume Determination using Radio-iron.*[3] Two useful radio-isotopes (p. 209) of iron are available: $^{59}Fe^*$, half-life 47 days, emits γ- and β-rays; $^{55}Fe^*$, half-life 5 years, emits X-rays. Ferric ammonium citrate equivalent to 1 mg. of Fe, with a *trace* of its iron as the radio-active isotope $^{55}Fe^*$ or $^{59}Fe^*$, is injected intravenously into a subject (called the donor) who should belong to Group O and be Rh negative. The radio-iron is taken up by the red bone marrow of the donor and is used in the manufacture of new red cells. A measured volume of the donor's blood is withdrawn and the radio-activity of a cell sample is determined; the donor's radio-active blood is then injected into a recipient whose cell volume is to be measured. The injected "tagged" (radio-active) red cells are thoroughly mixed up by the circulation with the recipient's cells. A specimen of the recipient's blood is withdrawn and the degree of dilution of radio-activity in a cell sample is determined; from this the degree of dilution of the injected red cells by the recipient's red cells and thus the recipient's red cell volume is determined. The most reliable expression for the ratio of red cells to plasma in the *entire* blood volume is

$$\frac{\text{red cell volume determined by radio-Fe}}{\text{plasma volume determined by radio-I-protein.}}$$

EXTRACELLULAR WATER.—Ideally this volume could be determined by measuring the degree of dilution of a non-toxic substance with these

[1] It is noteworthy that as a matter of fact 10–25% of the iodinated protein escapes from the plasma in the first hour (about the same rate at which the blue dye escapes).

[2] Gibson *et al., J. clin. Investig.*, 1946, *25*, 848.

[3] Gibson *et al., J. clin. Investig.*, 1946, *25*, 605, 616.

properties: (i) it can pass freely through the capillary endothelium and become uniformly distributed throughout the extracellular water; (ii) it cannot pass through the cell membranes; (iii) during the mixing period it is neither metabolized nor excreted. Radio-active Na [$^{24}Na^*$] or Cl [$^{38}Cl^*$] (given as NaCl) may be used; but even these ions penetrate to a small extent into certain tissue cells and so give values for the extracellular water which are too high. If the volume of the extracellular water and the plasma are known, the volume of interstitial water can be derived by subtraction.

INTRACELLULAR WATER.—To determine this volume, total body water must be determined by administering a substance which diffuses freely throughout the water of the body, *i.e.* it must be able to pass through all cell membranes (as well as through the capillary endothelium). The only known substance with the necessary properties is heavy water (D_2O), *i.e.* water in which the ordinary H atoms have been replaced by deuterium (D, heavy hydrogen, of atomic weight 2). Heavy water in the concentrations used in such experiments has the same properties as water and like water moves freely into and out of cells.[1] If total body water is known, on subtracting extracellular water the volume of the intracellular water is derived.

SOME ELEMENTARY CHEMICAL AND PHYSICO-CHEMICAL DATA REVIEWED [2]

It is essential to have a clear understanding of the following terms:

ELECTROLYTES. IONS.—All acids, bases and salts (*electrolytes*) in aqueous solution dissociate to a varying extent into electrically charged particles called *ions*. Thus, in the body fluids NaCl does not exist as the molecule, but as positively charged Na ions (Na^+) and negatively charged Cl ions (Cl'). Similarly, $NaHCO_3$ dissociates into Na^+ and HCO_3'. In an electric field all positively charged ions (H^+ and the so-called basic radicals, *e.g.* Na^+, K^+, NH_4^+, Ca^{++}, Mg^{++}) move to the negative pole or cathode, and are therefore called *cations*; all negatively charged ions (OH' and the so-called acid radicals Cl', HCO_3', HPO_4'', SO_4'') move to the positive pole or anode, and are therefore called *anions*. At the slightly alkaline reaction of the body fluids the proteins exist as salts (*i.e.* combined with bases), *e.g.* in plasma as Na Proteinate which ionizes into Na^+ and an enormous acidic anion, Proteinate$'$ (p. 134).

It is important to consider the behaviour of the *individual ions separately* (and not in terms of the salts from which they are derived) as the concentrations of any ion may vary independently of the others. Thus the concentration of Na^+ and Cl' may vary independently of one another. Certain of the *organic* solutes in the body fluids are un-ionized; *e.g.* urea and glucose exist as the molecules and are thus *non-electrolytes*.

MOLAR SOLUTION.—A molar solution contains the molecular weight of the substance expressed in grammes (g.) dissolved in 1 litre (L). Thus a molar solution of NaCl (atomic weight of Na is 23; of Cl is 35·5; molecular weight of NaCl is 58·5) contains 58·5 g. of NaCl per L.

[1] For technique see Schloerb *et al.*, *J. clin. Investig.*, 1950, *29*, 1296.

[2] Davson, *General Physiology*, London, 1951 (full treatment of physico-chemical aspects of physiology).

VALENCY.—A *monovalent* atom or radical (*e.g.* Na^+, K^+, NH_4^+; OH', Cl') combines with or displaces one atom of hydrogen; a *divalent* atom or radical (*e.g.* Ca^{++}, Mg^{++}; SO_4''), combines with or replaces two, and a *trivalent* three, atoms of hydrogen.

EQUIVALENT WEIGHT.—The *gram-equivalent weight* (loosely referred to as the "equivalent weight" or the "equivalent") of a reactive unit (atom, ion, molecule) is the weight in g. which combines with or displaces 1 g. of hydrogen. Thus the equivalent weight of monovalent units (Na^+, K^+, OH', Cl', NH_4^+, NH_3, etc.) is *equal* to their unit weight (atomic-, ionic-, molecular-weight); the equivalent weight of divalent units (Ca^{++}, Mg^{++}, SO_4'', H_2SO_4) is *half* their unit weight. The equivalent weight of any unit combines with or replaces the equivalent weight of any other appropriate unit. Thus 23 g. of Na^+, 39 g. of K^+, or 18 g. of NH_4^+ mutually replace one another in salts, or combine with 35·5 g. of Cl', 48 g. of SO_4'', or 61 g. of HCO_3'; the weights mentioned above are *chemically* equivalent (hence the term equivalent). One equivalent of A will combine with or replace one equivalent of B whatever the valencies or unit weights of A and B. By expressing the concentration of units in terms of equivalents we eliminate the factor of difference in the *weight* of the individual units and can compare the relative *number* of chemical combining units concerned. This device is made use of in expressing the concentration of H^+ and OH' ions when the acidity or alkalinity of a solution is being described. As the electrolyte concentrations in the body fluids are low it is customary to express them as *milliequivalents* (m.Eq.). 1 m.Eq. is one thousandth of a g. equivalent per litre; *i.e.* to convert equivalent into milliequivalents multiply by 1000. Thus a solution containing 23 g. of Na^+ per L=1 Eq. per L=1000 m.Eq. per L.

As ionic and electrolyte concentrations are more generally expressed as weight per volume, one must be able to convert concentrations expressed as mg. per 100 c.c. into concentrations expressed as m.Eq. per L. The following conversion table (Gamble) shows how the calculation is carried out.

The general formula used is: value in m.Eq./L=mg. per 100 c.c.×10×(V÷A) (where V=valency; A=atomic weight).

m.Eq./L

Na^+:	mg.	per	100	c.c.×10	÷23 (V=1)
K^+:	,,	,,	,,	,,	÷39 (V=1)
Ca^{++}:	,,	,,	,,	,,	×(2÷40) (V=2)
$Mg.^{++}$:	,,	,,	,,	,,	×(2÷24) (V=2)
Cl':	,,	,,	,,	,,	÷35·5 (V=1)
HPO_4'':	,,	,,	,,	,,	×(1·8 [1]÷31) (V=1·8) (expressed as P)
SO_4'':	,,	,,	,,	,,	×(2÷32) (V=2) (expressed as S)

It is customary to express the concentration of carbonic acid (H_2CO_3) and of bicarbonate ($BHCO_3$) in blood not as a mass (*i.e.* a weight) but as the

[1] In the plasma the inorganic phosphate is 20% in the form of BH_2PO_4 (containing one Eq. of base (B)) and 80% in the form of B_2HPO_4 (containing two Eq. of base).

BH_2PO_4 ionizes into B^+ and H_2PO_4' (univalent).
B_2HPO_4 ionizes into $2B^+$ and HPO_4'' (divalent).

Though the inorganic phosphate in the body fluids is always represented as the divalent form HPO_4'', the virtual valency is 1·8 *i.e.* (0·2×1)+(0·8×2)).

volume of CO_2 (measured at N.T.P.) which the H_2CO_3 or $BHCO_3$ would give off after acidification in a vacuum. These vols. of CO_2 per 100 c.c. blood are converted into m.Eq. per L by dividing by 2·22.

Thus 3 vols. per cent. of CO_2 as $H_2CO_3 = 3 \div 2{\cdot}22 = 1{\cdot}4$ m.Eq./L
50 vols. per cent. of CO_2 as $BHCO_3 = 50 \div 2{\cdot}22 = 22{\cdot}7$ m.Eq./L.

PHYSICAL FACTORS INVOLVED IN FLUID AND ELECTROLYTE EXCHANGE

Before considering in detail the movement of water and dissolved substances (solutes) between the various compartments of the body certain physical considerations must be recalled.

FILTRATION.—By filtration is meant the passage of water and dissolved substances through a membrane owing to a difference of hydrostatic pressure on the two sides of the membrane. This difference of pressure is termed the filtering force. The fluid which passes through is termed the filtrate, and consists of water and any dissolved substances to which the membrane is permeable. Filtration characteristically occurs in the capillaries where the capillary blood pressure drives a more or less protein-free plasma out of the blood vessels.

PROPERTIES OF MEMBRANES.—Membranes differ very much in permeability, *i.e.* in the size and kind of molecule which is allowed to pass through them. Certain artificial membranes only let through water, but are impermeable to all dissolved substances, *e.g.* the well-known copper ferrocyanide pot;[1] such membranes are called "semi-permeable."[2]

(i) The membranes of the tissue cells show the peculiar differential permeability fully discussed on p. 7.

(ii) The capillary endothelium (p. 6) is freely permeable to all the dissolved crystalloids (as well as water), but normally it is more or less completely impermeable to protein and similar large colloidal particles. In the glomeruli of the kidney proof is available that a fluid is filtered through the capillary wall which is identical in composition with plasma but minus its protein and fat. Damage to the capillary wall increases its permeability to protein; badly damaged capillaries allow whole plasma (with all its protein) to escape; in graver injuries the corpuscles pass out too.

DIFFUSION.—This is the movement of solutes across a membrane separating two fluids under the same hydrostatic pressure but containing different concentrations of the solute. Thus glucose, urea, Na^+, Cl' will diffuse from a region of high to a region of low concentration.

OSMOTIC PRESSURE. OSMOSIS.—These terms are difficult to define but are readily illustrated by reference to an experiment. If a strong glucose solution is placed inside a copper ferrocyanide pot (which is permeable to

[1] Copper ferrocyanide is produced when copper sulphate comes into contact with potassium ferrocyanide; the precipitate can be formed in the pores of a porous pot, which is a convenient support for it, by putting copper sulphate inside the pot and potassium ferrocyanide outside.

[2] This word is sometimes used to describe membranes (like those of the capillaries which are permeable to water and crystalloids, but hold back colloids. This should be avoided and the word should only be used in the sense given above. There are *no* true semi-permeable membranes in the body.

water but impermeable to glucose), and the pot is immersed in water, water passes from the exterior into the interior of the pot and the pressure within it rises to a certain extent. This rise of pressure is a measure of the osmotic pressure of the glucose solution;[1] it is as though the glucose molecules within the pot exerted a *suction action attracting water to themselves* with a certain force.

It should be remembered that the osmotic pressure of a solution depends solely on the *number of particles* (undissociated molecules and ions) in solution and not on their size. The molecular weight in grammes of an (undissociated) substance, dissolved in 1 litre of water, exerts an osmotic pressure of 22·4 atmospheres. For example: the molecular weight of glucose is 180; a solution of 180 g. of glucose in 1 litre of water exerts 22·4 atmospheres pressure; 0·1% solution of glucose (=1 g. in 1 litre, the concentration in plasma) exerts an osmotic pressure of $\frac{22 \cdot 4}{180}$ atmospheres=0·125 atmospheres =0·125×760 mm. Hg=95 mm. Hg.

When dealing with the osmotic pressure (o.p.) of solutions, the concentration of osmotically active particles is often best expressed as *osmoles* (or *milliosmoles*) per L. One osmole (equals 1000 milliosmoles) of a substance is that quantity which gives an o.p. of 22·4 atmospheres when present in 1 L of solution. Thus, for a solution of an *undissociated non-electrolyte*, the molar and osmolar values are identical; an *osmolar* solution (1 osmole/L) will contain 1 g.-mole of a non-electrolyte/L, *e.g.* 180 g. of glucose or 60 g. of urea per L. In the case of substances which completely dissociate into ions, *each ion* exerts the same o.p. as an undissociated molecule. Thus a molar solution of NaCl ($Na^+ + Cl'$) has an o.p. of 2×22·4 atmospheres, and is therefore 2 osmolar (2 osmole/L). If substances dissociate only partially, the appropriate multiplication factor must be found by experiment.

We refer to osmolar concentration (osmolarity) instead of to molar concentration when we wish to make clear the *additive* osmotic effect of individual ions (pp. 30–32).

Solutions with the same osmolarity are termed *isotonic* (or *isosmotic*). A solution termed "isotonic" (without further qualification) is one that is isotonic with normal plasma (though strictly speaking a single solution cannot be called isotonic without some other reference solution being named); similarly, *hypertonic* or *hypotonic* solutions have osmotic pressures respectively greater or less than normal plasma.

OSMOTIC PRESSURE OF PLASMA.—The *total* osmotic pressure of plasma is almost wholly due to its dissolved inorganic salts and other crystalloids, but to a slight extent it depends on the *plasma proteins*. The calculated osmotic pressure of the plasma is 6·5 atmospheres (*i.e.* about 5000 mm. Hg), but in the body only a minute fraction of this immense force is available for suction purposes, because under normal conditions the interstitial fluids (tissue fluids) have about the same crystalloid concentration as plasma, *i.e.* they are isosmotic (isotonic) with plasma; as the capillary membrane is freely permeable to crystalloids their concentrations on the two sides of membrane are usually the same.

[1] The osmotic pressure of a solution can be also determined indirectly by the *depression of the freezing-point*, which is proportional to the concentration of undissociated molecules and of ions of solute present.

IMPORTANCE OF CRYSTALLOID OSMOTIC PRESSURE AND ELECTROLYTE CONCENTRATION.—The crystalloid osmotic pressure of the body fluids and of the tissue cells is one of the most important "constants" of the internal environment; its normal value is maintained far more rigidly than is that of the volume of the body water. Quite minor variations in crystalloid osmotic pressure are very harmful, as is easily shown by studying the reactions of isolated organs like heart, skeletal muscles, and intestine in solutions which are hypotonic or hypertonic. Crystalloid osmotic pressure is almost entirely due to the *electrolytes* which form a sort of *fixed chemical framework* to the body fluids; the osmotic pressure of the organic non-electrolytes is quantitatively small and more variable. For proper functioning of the tissues the concentration of *each* individual ion must be accurately maintained as well as the total osmotic pressure of all the electrolytes.

Plasma Protein Osmotic Pressure.—The osmotic pressure of the *plasma proteins* (though very small) is of great functional importance, because the capillary endothelium is relatively impermeable to protein. It can be determined as follows: An osmometer is employed, which consists of a small glass bell from the top of which project two vertical pieces of tubing. Over the mouth of the bell is tied the membrane, such as cellophane, which is completely impermeable to protein (but is freely permeable to water and dissolved crystalloids); plasma is placed in the bell, and the apparatus is immersed in a solution having the same electrolyte concentration as plasma but free from protein, *e.g.* Ringer-Locke's solution. In the course of a few days it is found that the fluid in the projecting tubes rises to a height of about 350 mm., equivalent to about 25 mm. Hg. The proteins have exerted an osmotic pressure of 25 mm. Hg and have caused fluid to flow into the bell. As the membrane is permeable to the electrolytes in solution, they also diffuse through, so that the ultimate result is the addition of a certain volume of physiological saline to the plasma.

As the protein concentration in the interstitial fluid is negligibly small, the *plasma protein osmotic pressure helps to attract fluid into the blood vessels.* On the other hand, the capillary blood pressure is a filtering force tending to drive fluid into the interstitial spaces, and thus operates in the opposite direction to the osmotic pressure of the colloids.

The osmotic pressure of the plasma proteins depends mainly on the serum albumin and to a smaller extent on the serum globulin. The osmotic pressure exerted by a substance depends on the *number* of molecules in the solution; as albumin has a much smaller molecule than globulin, a given weight of albumin contains many more molecules than the same weight of globulin, and therefore exerts a higher osmotic pressure. 1 g. albumin per 100 c.c. exerts an o.p. of 6·0 mm. Hg and 1 g. globulin per 100 c.c. only 1·5 mm. In addition, the specific osmotic pressure of a protein (*i.e.* osmotic pressure in mm. Hg per g.) rises with increase in its concentration; conversely reduction of the plasma protein concentration lowers the osmotic pressure per gramme of the remaining plasma protein. In hypoproteinæmia, therefore, there is a double disadvantage; (i) the total protein concentration is reduced, and (ii) the specific osmotic pressure (mm. Hg per g.) is decreased. With the normal albumin/globulin ratio of 1·7/1·0, by far the greater part of the plasma protein osmotic pressure (about 80%) is due to albumin. By virtue of their osmotic pressure the plasma proteins tend to retain fluid in the

capillaries and so help to maintain the plasma volume (p. 81) and regulate the interchanges between the blood and the tissue spaces (p. 18); each g. of albumin which is added to the plasma sucks into, and retains in it about 20 c.c. of fluid. The plasma proteins influence filtration in the glomeruli of the kidney (p. 27). If the plasma protein concentration is decreased experimentally (by plasmapheresis), or by disease (*e.g.* Type 2 nephritis [nephrosis]) to below 3·5%, extensive œdema develops (p. 113).

EXCHANGE OF FLUID BETWEEN BLOOD AND TISSUE SPACES. LYMPH

Tissue Spaces and Lymphatics.[1]—The capillaries are a closed system of vessels; the blood only comes into actual contact with the cells of the tissues in the liver, where the endothelial lining of the vascular capillaries is deficient, and perhaps also in the spleen. The cells everywhere else are bathed by tissue fluid (interstitial fluid), which acts as an intermediary, supplying nutritive materials and receiving the products of metabolic activity.

FIG. 3.—Arrangement of Lymphatic Capillaries. Reconstruction of small portion of skin of leg of fœtus (×70). (After Drinker and Kampmeir, *Amer. Heart. J.*, 1939, *18*, 397.)

The extreme wealth of the lymphatic network in the skin is well shown. Note the numerous valves. The subcutaneous plexus drains into the more slender and more regularly disposed and fainter) deeper lymphatic channels.

The lymphatic system consists of the lymphatic capillaries and the larger vessels which unite to form the right lymphatic duct and the thoracic duct. The lymphatics are a closed system of vessels where they ramify in the tissues (Fig. 6). A distinction should therefore be drawn between the fluid in the tissue spaces and that in the lymphatic channels, to which the term lymph should be restricted (Fig. 3).

The large serous cavities like the peritoneum and pericardium can be regarded as specialized tissue spaces. It is doubtful whether any stomata connect the serous cavities with the lymphatics.

COMPOSITION OF LYMPH.—Lymph from the lymphatic ducts has a higher protein content than the fluid in the tissue spaces and may contain a few lymphocytes. Thoracic duct lymph contains about 4% of protein; it is rich in fat after a fatty meal (p. 865). The lymph coming from the limbs, intestine, and liver (in animals), contains 2, 4, and 6% of protein respectively. In an isolated observation, the lymph, from the ankle lymphatics in a woman during walking, contained 0·5% of protein. As lymph is derived from tissue fluid, its comparatively high protein content is not easily accounted for.

[1] Drinker and Yoffey, *Lymphatics, Lymph and Tissue Fluid*, Cambridge, Mass., 1941. Symposium on Lymph, *Ann. N.Y. Acad. Sci.*, 1946, *46*, 679–882.

Exchange of Fluid between Plasma and Tissue (Interstitial) Fluid.—This takes place owing to:

(i) *Differences in the crystalloid osmotic pressure of the fluid in the two compartments.* Such exchanges result from water drinking (p. 55), water deprivation (p. 66), salt excess (p. 62), salt deprivation (p. 64), and injection of hypotonic or hypertonic saline (pp. 126, 127); they are fully discussed at the pages mentioned. During *tissue activity* osmotically active metabolites are released into the interstitial fluid causing water to flow out from the plasma (p. 20).

(ii) *Differences between (a) capillary blood pressure and tissue fluid pressure, (b) plasma protein and tissue fluid protein osmotic pressure.* The exchanges occurring in resting and active tissues are considered below.

CAPILLARY BLOOD PRESSURE.—This can be measured in man by introducing a very fine cannula filled with saline (connected with a hand pump and a manometer) into a skin capillary. If the pressure in the capillary is higher than in the glass tube, blood runs into the saline, and vice versa. The pressure is then suitably raised or lowered and tested on other capillary loops. The following results have been obtained[1] in the capillaries at the root of the nail, with the subject in the recumbent position and the finger held at the level of the manubrium sterni. There is a considerable pressure gradient in a capillary loop, the pressure being higher at the arterial than at the venous end of the capillary (Fig. 4). The *average* values are: arterial end, 32 mm. Hg; summit, 20 mm. Hg; venous end, 12 mm. Hg. At the arterial end there is a difference of 5–10 mm. between the pressures during systole and diastole (*i.e.* the flow is pulsatile); but at the venous end the pulse pressure is negligible (the flow is uniform). These results indicate that the capillaries constitute to some extent a part of the peripheral resistance (cf. p. 302).

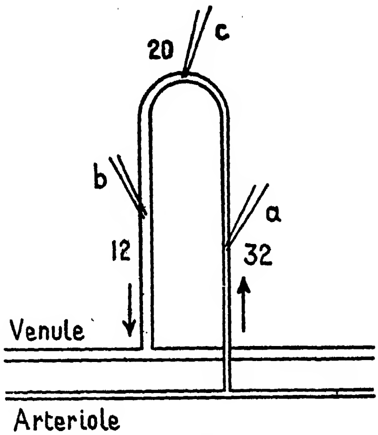


FIG. 4.—Capillary Blood Pressure. *a*—Arteriolar capillary; *c*=Summit of loop; *b*=Venous capillary. The figures (32, 20, 12) indicate average capillary pressure in mm. Hg in each part of the capillary. (Landis *Heart*, 1930.)

Relation to Heart Level.—When the finger is raised above the level of the manubrium sterni (*i.e.* above heart level), there is practically no change in the venous capillary pressure, and only a slight fall in the arterial capillary pressure. If the finger is lowered below heart level, the capillary pressures rise by almost the full theoretical hydrostatic requirements (Fig. 5).

Relation to Venous Pressure.—If the venous outflow is obstructed, the capillary pressure rises and attains a value about 10 mm. Hg higher than the venous pressure (cf. p. 322).

Effects of Arteriolar Dilatation.—This can be conveniently studied by

[1] Landis, *Heart*, 1930, *15*, 209.

heating the skin and so dilating the local arterioles. Both arterial and venous capillary pressures markedly rise (*e.g.* from 40 to 60, and 10 to 40 mm. Hg respectively), and capillary pulsation with each heart beat is obvious. The rise is due to the general arterial pressure being transmitted more directly to the capillaries owing to the lessened resistance in the wider skin arterioles.

Internal Organs.—Less is known about capillary pressure in internal organs. The pressure in the glomeruli of the kidney may be nearly as high as the diastolic blood pressure (p. 25). In the intestinal capillaries the pressure must be higher than in the portal vein, which in man is 10–20 mm. Hg. In the liver, capillary pressure may be very low indeed.

Plasma-Tissue Fluid Exchanges in Resting Tissues.—At the *arterial*

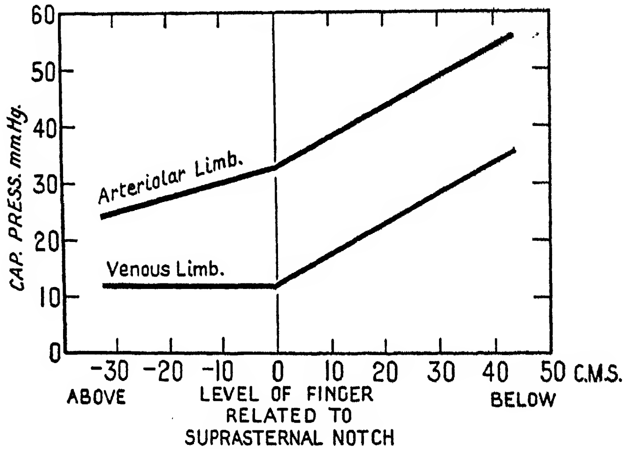


FIG. 5.—Effect of Gravity on Capillary Blood Pressure.

Vertical line 0=Finger at heart level. Levels above are recorded to the left as minus; below, to the right as plus. (Landis, *Heart*, 1930.)

end of a capillary the balance of the outward-driving (capillary pressure) and inward-pulling (protein osmotic pressure) forces is such that small quantities of fluid and dissolved substances leave the vessels to enter the tissue spaces. The average pressure in an arterial capillary (in the skin of man at heart level) is 32 mm. Hg; the protein osmotic pressure in the plasma is 25 mm. This gives an outward-driving force of 7 mm. At the venous end of the capillary the pressure is lower, approximately 12 mm.; this is less than the protein osmotic pressure which has simultaneously risen slightly owing to previous loss of fluid from the capillary. Consequently an approximately equal volume of fluid is reabsorbed (Fig. 6). This plasma-tissue fluid interchange takes place on an enormous scale; a volume of fluid equal to that of the whole plasma escapes from the entire capillary bed into the tissue spaces in less than one minute.

In the dependent parts, where the hydrostatic pressure of the column of blood may be considerable, the capillary pressure, even at the venous

end, may exceed the protein osmotic pressure. When the arm is allowed to hang limply, it may swell from filtration of fluid outwards into the tissue spaces; the venous blood leaving the part is correspondingly more concentrated.

Plasma-Tissue Fluid Exchanges in Active Tissues.—During tissue activity larger quantities of fluid pass out of the capillaries into the tissue spaces. In the case of a *secreting gland* (*e.g.* salivary gland) most of this fluid is then *actively transferred* by the gland cells from the tissue spaces into the lumen of the gland duct. This active process is called *secretion* and

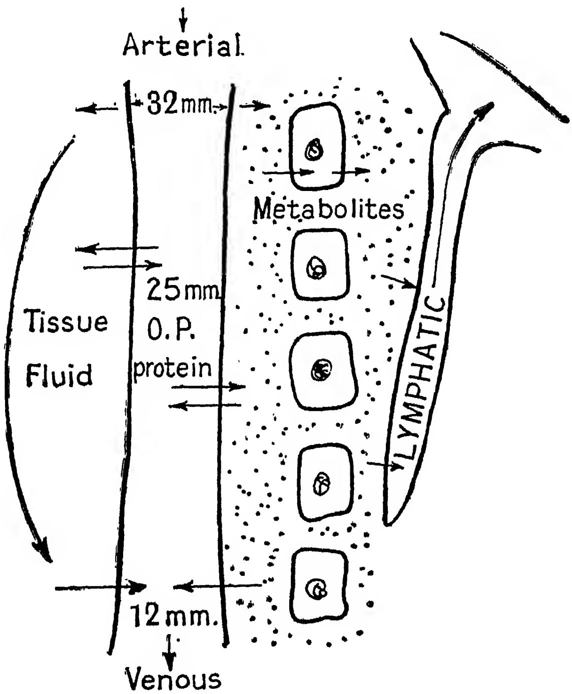


FIG. 6.—Fluid Interchanges between Plasma, Tissue Spaces, and Lymphatics. (Wright, *Proc. roy. Soc. Med.*, 1939.)

unlike purely physical processes like osmosis or filtration it involves increased metabolism (increased O_2 consumption and CO_2 formation, increased heat output). Some of the fluid passes by unknown means into the lymphatics and so leaves the active organ; some accumulates in the tissue spaces.

In the case of *muscle*, tissue fluid formation is on a smaller scale; the excess fluid can neither escape in the lymphatics or remain in the muscle. In muscular activity the weight of the muscle may increase by 20%; this swelling of the muscle may be in part responsible for the stiffness which follows severe exertion. The greatly enhanced rate of tissue fluid formation supplies the active organ with more nutrients and other materials (including

water) necessary for the formation of the secretion or for the carrying out of the specific function of the organ. But it should be remembered that the nutrients could have been provided by simple diffusion owing to a difference of concentration of the nutrient on the two sides, without any escape of fluid, as is the case when O_2 is supplied to an active tissue. During tissue activity when, possibly, no fluid passes back into the blood, the metabolites which are formed in the tissues gradually diffuse away into the blood owing to difference of concentration and are thus removed.

The factors concerned in this greatly enhanced tissue fluid formation are: (i) changes in the circulation and (ii) release of metabolites.

VASCULAR CHANGES.—(1) *Blood flow.*—The blood flow through an active organ is greatly increased owing (i) to local arteriolar dilatation; this is mainly due to the liberation in the organ of vasodilator products of activity ("metabolites," *infra*); sometimes it is due to the action of vasodilator nerves. Thus the chorda tympani nerve is not only secretory to the submaxillary gland but is also specifically vasodilator (p. 316); (ii) to diversion of a larger proportion of the cardiac output to the active region. In muscular exercise not only is the cardiac output greatly increased but most of it is sent to the active muscles (p. 433). The magnitude of the blood supply sets an upper limit to the volume of tissue fluid which can be formed.

(2) *Increased Arteriolar Pressure.*—Owing to arteriolar dilatation there is reduced frictional resistance and the arterial pressure is transferred from the proximal to the distal end of the arteriole with less loss of pressure.

(3) *Capillary Changes.*—Owing to the action of metabolites and local rise of temperature (p. 320) the capillaries in the active region become widely dilated and many capillaries previously closed open up; the capillary membrane becomes thinner and more permeable permitting more rapid transfers across it. As the blood pressure is transmitted with less loss through the arterioles, the capillary blood pressure rises; as the capillaries are also dilated there is less local frictional resistance; consequently the pressure at the venous end of the capillary corresponds more closely to that at the arterial end. It is quite likely that throughout the length of the capillary the blood pressure may exceed the opposing protein osmotic pressure; if this is so, this factor alone would lead to a continuous outflow of fluid with no possibility of its return into the capillaries.

RÔLE OF METABOLITES.—During activity, large non-osmotic molecules of the resting organ break down into numerous small molecules, which pass into the tissue spaces. These metabolites diffuse *slowly* into the blood and remain in considerable concentration in the tissue spaces for a long period. Such metabolites as are osmotically active attract fluid from the blood and retain it. The dissolved crystalloids of the blood, which can pass very readily through membranes, diffuse through into the tissue spaces at the same time as the water is being "pulled" there, *i.e.* protein-free plasma passes out. Very little is known about the metabolic changes taking place in active glands; in the case of muscle, however, a substantial rise of osmotic pressure undoubtedly does take place during activity (p. 430, footnote).

It is worth emphasizing that exudation of fluid from the blood vessels is synchronous with, and due to, the *functional activity of the organ*; for the products of activity (metabolites), not only raise the osmotic pressure in

the tissue spaces but are largely responsible for the arteriolar and capillary dilatation as well (pp. 317, 320).

REMOVAL OF TISSUE FLUID.—During the period of activity, the tissue fluid which is formed is (i) retained in the tissue spaces or (ii) escapes into the lymphatics (this may be the purpose of the lymphatics in the normal body in the absence of infection); (iii) in the case of *glands* it is mainly poured out in the external secretion. When the tissue comes finally to rest the metabolites either diffuse into the blood or are oxidized, and thus the osmotic pressure of the tissue spaces falls. The arterioles and capillaries become narrowed again. The excess fluid in the tissue spaces is absorbed mainly into the blood vessels, and possibly also into the lymphatics. As the lymphatics are a closed system of channels peripherally, no fluid can enter them directly, but must first pass through the endothelial lining. We do not know how this transfer is accomplished.

In the above discussion attention has been directed mainly to the exchanges between plasma and tissue (interstitial) fluid. The exchanges between extracellular and intracellular fluid are considered on pp. 55–68.

PLASMA-TISSUE FLUID EXCHANGES AFTER INJECTION OF HYPOTONIC OR HYPERTONIC SALINE.—See pp. 126, 127.

TISSUE FLUID AND PLASMA VOLUME REGULATION.—See p. 80.

REMOVAL OF FLUID FROM SEROUS CAVITIES AND SUBCUTANEOUS TISSUES. —If isotonic saline is introduced into a serous cavity it is absorbed very rapidly. In the case of the pleural cavity, ligature of the thoracic and right lymphatic ducts does not affect the rate of absorption of saline, proving that the fluid is passing chiefly into the blood. If a dyestuff, *e.g.* indigo blue dissolved in saline, is injected into the pleural cavity, it appears in the urine in 5 minutes, and in the thoracic duct after only half an hour has elapsed. Absorption into the lymphatics therefore does occur, but is relatively unimportant and slow.

Blood which has effused into a serous cavity may also be slowly absorbed. It is assumed that the proteins of the effusion undergo autolysis and subsequent removal, or are ingested by phagocytic cells. The protein-free plasma is then absorbed.

FUNCTIONS AND STRUCTURE OF THE KIDNEY. SECRETION OF URINE[1]

Functions of the Kidney.—1. By *secreting the urine* the kidney carries out many essential functions.

(1) The kidney regulates the *volume* of the body fluids; more specifically it maintains the constancy of the volume of the extracellular fluids, with special emphasis on the volume of the *plasma*. Water is, of course, lost by other channels; in the case of the skin, the volume of water lost as sweat

[1] Gamble, *Extracellular Fluid*, Harvard University Press, 1949. Homer Smith, *Kidney: Structure and Function*, London and New York, 1951; *Studies in Physiology of the Kidney* (Porter Lectures), University of Kansas, 1939. *Bull. N.Y. Acad. Med.*, 1947, *23*, 177. *Technique of Renal Clearance*, see Smith *et al.*, *J. clin. Investig.*, 1938, *17*, 263; Goldring *et al.*, *ibid.*, 1940, *19*, 739. For review of renal function in early life, see McCance, *Physiol. Rev.*, 1948, *28*, 331.

is adjusted in relation to the "needs" of body temperature. The volume of the urine, however, is specifically adjusted to the "needs" of *water balance.*

(2) (i) The kidney is the only significant route by which inorganic ions, both cations (*e.g.* Na^+, K^+) or anions (*e.g.* Cl', phosphate, SO_4'') are eliminated from the body. By adjusting the relative excretion of water and of the individual ions, both the *total crystalloid osmotic pressure* and the *ionic pattern* (*i.e.* the concentration of individual ions, notably Na^+, Cl', HCO_3') of the extracellular fluids, including the plasma, are kept within normal limits. If there is a "conflict" between the regulation of the volume of the extracellular fluids and the regulation of the total crystalloid osmotic pressure, the latter takes precedence over the former, *i.e.* crystalloid osmotic pressure is maintained in spite of deviations in the volume of body water.

(ii) The kidney also helps to preserve the appropriate plasma concentration of crystalloid *non*-electrolytes, of which *glucose* and *urea* are the most important.

(3) (i) Among the ions, but so important as to demand special mention, are the H^+ and OH' ions, the relative concentration of which determines the acidity or alkalinity of the blood and body fluids generally. Regulation of the H^+ ion concentration of the plasma depends on (*a*) physico-chemical reactions (p. 91); (*b*) vital reactions (p. 94). One of the vital reactions, *i.e.* the adjustment in the pulmonary ventilation, helps to preserve the constancy of the H^+ ion concentration by varying the volume of CO_2 which is eliminated; but the lungs can only deal with a *volatile* substance like CO_2. The elimination of appropriate amounts of *non-volatile* acid and basic radicals can be carried out *only* in the urine. Electrolytes must be excreted in the urine in such amounts as to maintain not only the correct *total* crystalloid osmotic pressure of the body fluids as mentioned in (2) above, but also the proper level of H^+ ion concentration. Sometimes the "needs" of H^+ ion concentration may be sacrificed in the "interests" of total crystalloid osmotic pressure.

(ii) The kidney also preserves H^+ ion concentration by *manufacturing* NH_3 which neutralizes the acid radicals of the blood, the resulting NH_4 salt being excreted in the urine.

(4) The kidney is the only route of elimination of the *waste* products, especially the nitrogen- and sulphur-containing substances derived from the metabolism of ingested protein or resulting from normal or abnormal cellular metabolism. The kidney likewise eliminates *toxic* substances (*e.g.* drugs) which have entered the body.

From what has been said above it is clear that the kidney is the outstanding "guardian," both with respect to sensitivity and scope, of the constancy of the milieu intérieur. *Failure* of renal activity results in many changes: alterations in the total crystalloid osmotic pressure or the concentration of individual ions (including H^+ ions) in the plasma and interstitial fluids; alterations in the concentration in the body fluids of useful non-electrolytes like glucose, or of waste products like urea; alterations in the volume of the plasma and of the total extracellular fluid. Increased permeability of the membranes in Bowman's capsule may lead to loss in the urine of plasma protein or even of blood.

2. The kidney also has certain functions which are *unrelated to the secretion of urine.*

(1) The renal cortex forms a proteolytic enzyme known as *renin* which is discharged into the general circulation; it acts on a plasma globulin called *hypertensinogen* to form a polypeptide called *hypertensin* which powerfully constricts peripheral blood vessels giving rise to hypertension. The stimulus to the discharge of renin is a decrease in renal artery pressure or renal blood-flow. Renin is probably secreted when the blood pressure falls; by increasing the peripheral resistance it may help to maintain normal arterial blood pressure. In conditions of *renal ischæmia* excess renin is liberated giving rise to marked sustained hypertension (so-called *ischæmic hypertension* (cf. p. 345)). As renin is an internal secretion, the kidney may, somewhat surprisingly, also be regarded as an endocrine organ.

(2) The renal tubular epithelium carries out certain *chemical transformations:*

(i) It forms NH_3 (p. 97).

(ii) It hydrolyzes organic hexose-phosphate by means of the enzyme phosphatase to liberate inorganic phosphate (p. 1002); some of the phosphate eliminated in the urine may come from this source and not directly from the plasma phosphate. Urinary phosphate may also be derived from the organic phosphate in tissue cells (p. 95).

(iii) It carries out detoxicating reactions, *e.g.* it combines the benzoic acid of the blood with glycine (available locally) to form *hippuric acid* (p. 830).

Structure of the Kidney.—The functional unit of the kidney is the *nephron* which consists of Bowman's capsule and the renal tubule proper (Fig. 7). Each human kidney contains one million nephrons which drain via the collecting tubules into the renal pelvis and thence into the ureter.

(1) *Bowman's capsule* is invaginated by a tuft of capillary vessels, the *glomerulus*; capsule and glomerulus together constitute the *Malpighian* (*renal*) corpuscle. The capsule is lined by a very thin epithelium. The structure of the Malpighian corpuscle is well adapted to its function which is to filter out a cell-free fluid from the blood identical in composition with the plasma, except for the proteins (and other colloids) which do not pass through.

(2) The *renal tubule* consists of three segments; the histological differences to be described are related to functional differentiation.

(i) The *proximal* tubule (including the first convoluted tubule) is lined by a cubical epithelium; the free external border of the cells contains many mitochondria arranged in parallel vertical columns producing a striated appearance (brush border).

(ii) *Thin* segment (of the loop of Henle): the cells are flattened and contain clear protoplasm and the lumen is narrower than anywhere else in the tubule.

(iii) In the *distal* tubule (including the second convoluted tubule) the lining cells are cuboidal and contain granular basal striations but no definite brush border.

(3) The *collecting* tubules are lined by clear cubical cells; they lead into the larger ducts of Bellini (lined by clear columnar cells) which open at the apex of the pyramids.

The loops of Henle, the collecting tubules, and the ducts of Bellini lie

in the medulla of the kidney; the other parts of the tubule are found in the cortex.

The two kidneys in man weigh 300 g. or 0·45% of the total body weight;

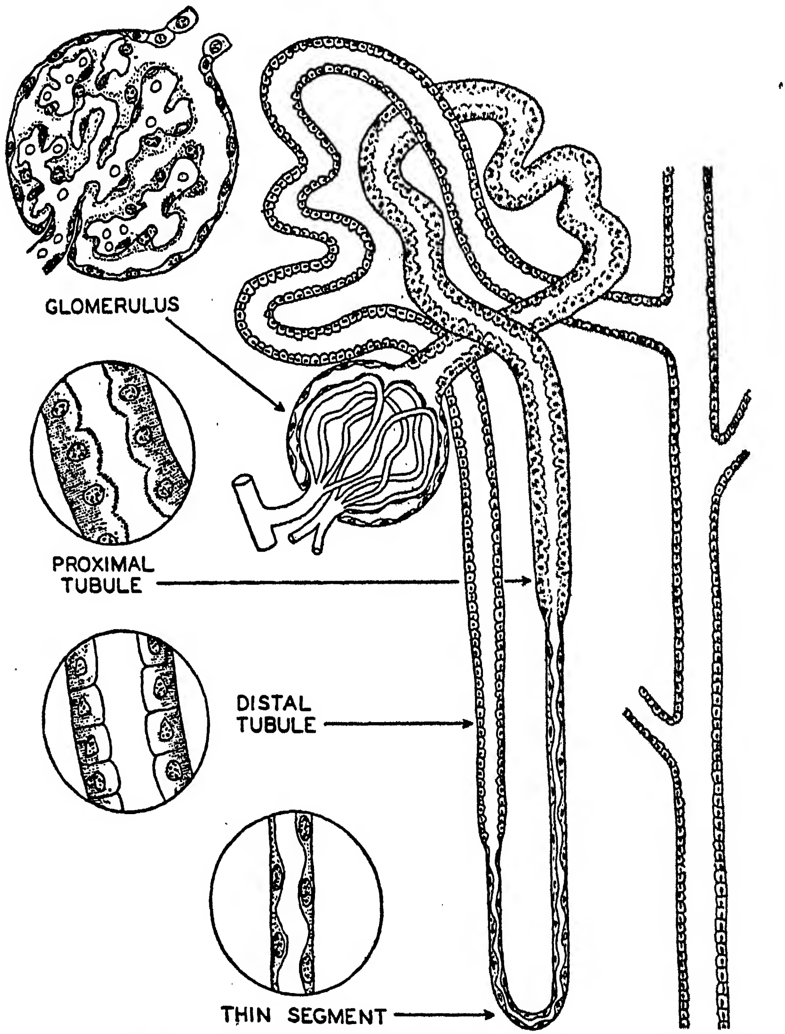


FIG. 7.—Arrangement and Structure of the Nephron. (Homer Smith, *Physiology of the Kidney*, 1937.)

the length of the human renal tubule is about 3 cm.; the diameter is 20–60 μ. As there are about 2 million nephrons, their total length is about 40 miles.

Renal Circulation.[1]—The blood flow through the kidney (as through the

[1] Trueta *et al.*, *Renal Circulation*, Oxford, 1947.

other organs) depends on the cardiac output, the general arterial blood pressure level, and the calibre of the local arterioles and capillaries. The distinctive features of the renal circulation are :

(i) The *enormous size of the renal blood flow*, about 1300 c.c. per minute in man. Although the kidneys are only 0·45% of the body weight, they receive about 25% of the resting cardiac output. The renal blood flow is not related to its respiratory needs but to its function of forming urine.

(ii) The *double capillary network*: (*a*) The glomeruli in the Malpighian corpuscles ; (*b*) the capillaries on the surface of the renal tubules.

(iii) The presence of a *shunt* mechanism which enables the blood flow to the renal cortex to be reduced, the blood then being diverted to the renal medulla (p. 26).

Blood Supply of Renal Cortex. The renal artery divides into branches which run between the pyramids to the " boundary zone " between the cortex and the medulla, where they are united by a series of *arterial arches* with their convexity directed outwards. From these arches, parallel *cortical* or *interlobular* arteries arise which run towards the surface of the kidney (Fig. 9). At intervals they give off the *afferent* vessels each of which breaks up into a capillary tuft, the glomerulus, whence emerges the *efferent* vessel which breaks up into a capillary net over the surface of the convoluted tubules ; the blood drains into venules and so enters the interlobular veins. These pass into the venous arches which lie in the concavity of the arterial arches. The efferent vessels of these glomeruli are narrower than the afferent vessels.

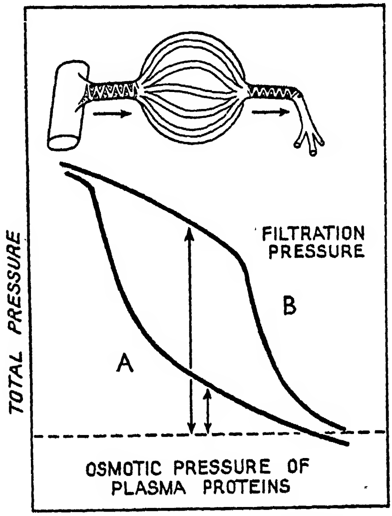


Fig. 8.—Glomerular Pressure as influenced by Relative Calibre of Afferent and Efferent Glomerular Vessels. (Homer Smith, *Physiology of the Kidney*, 1937.)

If the afferent vessel is constricted the glomerular pressure is markedly diminished (A) ; if the efferent vessel is constricted the pressure is raised in the glomerulus (B) and falls steeply on reaching the tubular capillaries. Height of arrows = net filtering force (p. 27).

Glomerular Blood Pressure.—The level of blood pressure in the glomerular capillaries though of great functional importance, has not been measured directly in mammals ; indirect evidence suggests that it is normally higher than the pressure in capillaries elsewhere and may sometimes equal the diastolic blood pressure. Glomerular pressure depends not only on the *general level of blood pressure* but also on *local* conditions within the kidney. With constant renal artery pressure, constriction of the *afferent* vessels of the glomeruli interposes an additional frictional resistance with resulting fall of the glomerular pressure (Fig. 8, A) ; conversely, constriction of the *efferent* glomerular vessels dams up the blood proximally so that glomerular pressure rises and approaches that in the renal arteries (Fig. 8, B). A rise in the filtration fraction (p. 27) in man is interpreted as representing raised glomerular

pressure. As explained below, in conditions of cortical ischæmia the pressure in the glomeruli and the blood flow through them is markedly reduced.

Blood Supply of the Medulla.—The glomeruli related to afferent vessels which spring from the interlobular artery close to its origin from the arterial arch (or occasionally from the arch itself), are known as *juxta-medullary glomeruli*, and have special characteristics (Fig. 9). Their efferent vessels are *wide* and may even exceed the afferent vessels in calibre. Each efferent vessel breaks up into a number of parallel "arterial" vessels, *vasa recta*, which run as a bundle into the *medulla* for varying distances before taking a hairpin bend or giving rise to a Y-branch, to return as "venous" vessels to

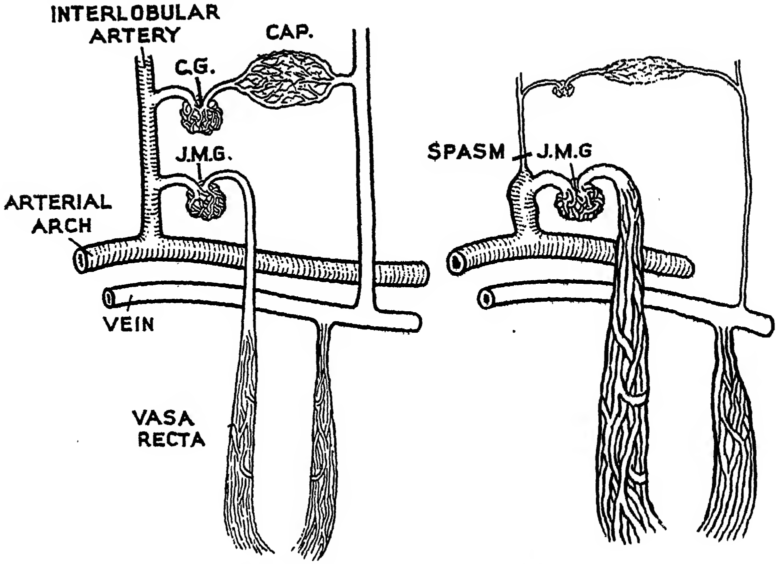


Fig. 9.—Diagram of Renal Circulation: Renal Shunts.

Left-hand Fig.: normal circulation.
C.G.=cortical glomerulus; Cap.=capillaries round convoluted tubules; J.M.G.=juxta-medullary glomerulus. Its efferent vessel breaks up into vasa recta which pass into the medulla.
Right-hand Fig.: spasm of interlobular artery just distal to juxta-medullary glomerulus. Note resulting cortical ischæmia and shunting of blood via juxta-medullary glomeruli into the medulla.

the venous arches. The vasa recta give off fine capillaries to the adjacent tubular tissue in the medulla. The vasa recta have a thin endothelium, like capillaries, but they are, of course, of *great length* and of much *wider calibre* than capillaries. Vasa recta may also arise directly from the arterial arches.

Renal Shunts. Cortical Ischæmia.—In a variety of experimental conditions the *interlobular arteries just beyond the point of origin of the juxta-medullary glomeruli* undergo *constriction*; as a result the blood flow to the cortex, and thus to the cortical glomeruli and related convoluted tubules, is almost completely cut off. The flow of urine is decreased or stopped. If the calibre of the other renal vessels and the general blood pressure are unaltered, the *total* renal blood flow is unaffected; but the blood which is excluded from the cortex is diverted via the juxta-medullary glomeruli and the vasa recta into the medulla (Fig. 9). The cortex is then pale and the medulla deeply

engorged with blood; as the diverted blood flow has not come into contact with the main mass of the nephrons and has, therefore, not given up its oxygen there, the venous blood is more arterial in colour than normally. If the cortical ischæmia is accompanied by constriction of the larger renal vessels or by a diminished renal blood flow from other causes the "shunt" will transfer less blood through the medullary vasa recta.

Localized *cortical ischæmia* with diversion of the blood flow to the medulla has been demonstrated in *rabbits* (i) during stimulation of the peripheral end of the splanchnic or renal nerves; (ii) reflexly from nocuous stimulation, *e.g.* crushing a limb; (iii) in rapid hæmorrhage. It is claimed that cortical ischæmia may be a factor in producing the anuria observed clinically (i) in emotional states, (ii) after operations on the urinary tract, (iii) after mismatched blood transfusions (cf. p. 182).[1] Prolonged severe cortical ischæmia may give rise to cortical necrosis.

Nerve Supply.—The sympathetic nerve supply of the kidney (p. 711) does not act directly on the nephrons; it modifies renal activity only by altering the local blood flow.

Mechanism of Renal Secretion.—An account will first be given of the probable course of events in the mammalian kidney; quantitative data quoted apply to man unless otherwise stated. The evidence for the various statements made will be set out subsequently (p. 31).

Glomerular Filtration.—About 1300 c.c. of blood (700 c.c. of plasma, 600 c.c. of cells) flow through the kidneys per minute; in the glomeruli, where the blood pressure is comparatively high, 120 c.c. of fluid are filtered off per minute from the 700 c.c. of plasma and passed into Bowman's capsule. The glomerular filtrate is identical in composition with the plasma but contains no protein (or other colloids). The ratio of the volume of glomerular filtrate/plasma flow is called the *filtration fraction*; in the example given it is 120/700=0·17. In 24 hours the total volume of glomerular filtrate is 170 litres.

The mechanism of glomerular filtration is like that responsible for tissue-fluid formation (p. 18). The glomerular blood pressure (say 75 mm. Hg) constitutes a filtering force driving fluid out of the blood vessels into Bowman's capsule; it is opposed by the osmotic pressure of the plasma proteins (=25 mm. Hg) which tends to hold water in the blood vessels. The *net filtering force* is the glomerular blood pressure minus the plasma protein osmotic pressure (*e.g.* 75—25=50 mm. Hg). The volume of glomerular filtrate depends on several factors: (i) directly on the magnitude of the net filtering force, *i.e.* it is increased by a rise in glomerular pressure or by a fall in plasma protein osmotic pressure; (ii) within certain limits, at constant filtering force, the volume of glomerular filtrate varies directly with the renal blood flow. (iii) The hydrostatic pressure in Bowman's capsule is normally zero (*i.e.* atmospheric pressure). Should it rise, *e.g.* as a result of obstruction to the outflow of urine, it resists the passage of fluid from the glomeruli; marked capsular distension from chronic ureteric obstruction may even lead to partial obliteration of the glomeruli.

In disease the renal blood flow and the number of functioning glomeruli may be markedly reduced; there is a close correlation between such reduction and the degree of resulting renal inefficiency.

[1] For critical review see Homer Smith *et al.*, *Amer. J. Med.*, 1950, *9*, 216.

The *permeability* of the filtering membranes (*i.e.* the glomerular endothelium plus the capsular epithelium) is increased in many abnormal conditions, *e.g.* by an inadequate blood supply (as in circulatory failure), anoxia, or by various toxic agents including certain drugs and bacterial poisons. The normal epithelia retain serum globulin (mol. wt.[1] 170,000) and serum albumin (mol. wt. 70,000), but they let through substances of smaller molecular weight, *e.g.* injected hæmoglobin (mol. wt. 68,000), egg albumin, Bence-Jones' protein or gelatin (mol. wt. 35,000). In renal disease, the membrane permeability usually increases: serum albumin is the first to escape and appears in the urine (*albuminuria*); more rarely and later serum globulin may also pass through. With graver injury to the Malpighian corpuscles both red and white cells may pass through or obvious *hæmorrhage* may take place. An abnormal state of the renal capillaries is probably responsible for the hæmaturia of scurvy (p. 1042) and purpura (p. 157).

Reabsorption in the Tubules.—The crystalloid composition of the glomerular filtrate is identical with that of the plasma. If the volume of glomerular filtrate is also known, the absolute amount of each plasma constituent filtered out (say in 24 hours) can be readily calculated. The urine for the same period can be collected and analysed. A comparison of glomerular filtrate and urine demonstrates that the glomerular filtrate is profoundly modified during its passage down the renal tubules (cf. Table, *infra*, which should be carefully studied).

Constituent	*Concentration in 100 c.c. of Plasma*	*Total in 170 L of Glomerular Filtrate*	*Total excreted in 24 hr. Urine*	*Total reabsorbed in Tubules*	*Concentration in mg. per 100 c.c. of Urine*	*Relative concentration Urine/Plasma*
Water .	..	170 L	1·5 L	168·5 L	..	..
Proteins, Colloids, Fats	8 g.	none	none	none	none	..
Glucose	100 mg.	170 g.	none	170 g.	none	..
HCO_3'	150	255	none	255	none	..
Na^+	330	560	5 g.	555	350	1
Cl'	365	620	9	611	600	2
Ca^{++}	10	17	0·2	16·8	15	1·5
K^+	17	29	2·2	26·8	150	9
Phosphate (as P)	3	5·1	1·2	3·9	150	50
Uric Acid	2	3·4	0·75	2·65	50	25
Urea	30	51	30	21	2000	70
SO_4'' (as S)	2	3·4	2·7	0·7	180	90
Creatinine		(data excluded owing to uncertainty.)				

Of 170 litres of water filtered out in the glomeruli, 168·5 litres are reabsorbed. Glucose is reabsorbed completely and bicarbonate at the normal *acid* reaction of the urine almost completely (but when an *alkaline* urine is passed large amounts of bicarbonate may be excreted (p. 96)). Na^+, Cl', and Ca^{++} are reabsorbed very extensively (only 1–2% of the amount filtered

[1] Mol. wt.=molecular weight.

is excreted); in the case of K^+, 8% of the amount filtered is excreted; in the case of uric acid this fraction is about 20%. In the case of phosphate and urea the fractions excreted are 25% and 60% respectively. There is controversy about the concentration of SO_4'' and creatinine in the plasma; if the data set out in the Table are reliable, about 80% of the SO_4'' filtered is excreted.

Substances which are absorbed completely or very extensively are called *threshold* substances; they are valuable plasma constituents which are absorbed in amounts sufficient to maintain their optimum concentration in the plasma. Substances which escape extensively and which are known to be waste substances (*e.g.* urea, sulphate, creatinine, and to a less extent uric acid) are called *no-threshold* (or low threshold) substances. As will be explained below the valuable threshold substances are probably *actively* reabsorbed; the fraction of the low threshold waste substances which returns to the blood does so by a *passive* process called *back-diffusion* (p. 31).

The last column in the Table compares the concentration of the constituents in urine and plasma. The threshold substances, because they are largely passed back into the blood are concentrated to a small extent in the urine. The no-threshold substances which return to the blood on a smaller scale are concentrated to a correspondingly higher degree.

It is quite clear from the above data that the fundamental process which takes place in the renal tubules is *reabsorption, i.e.* the transfer of water and many dissolved substances from the lumen of the tubules back into the blood. In the case of water and the threshold substances this process of absorption is *selective* and is *varied according to bodily needs.* It is likely that specific mechanisms are concerned with the reabsorption of each substance or group of substances. For example, the reabsorption of water is largely dependent on the action of the post-pituitary *antidiuretic hormone* (p. 49); adrenal corticoids increase the reabsorption of Na^+ and Cl' ions (p. 945); the reabsorption of glucose is arrested by *phloridzin* (p. 32). Experimentally, if the tubular epithelium is *paralysed by cyanide* or by *low temperatures* all active reabsorption ceases, proving that it is a *vital* process depending on the specific properties of the living epithelial cells (cf. p. 9).

Urinary Specific Gravity and Osmolarity.—The Table below shows the approximate relationship between the concentration of total solutes in the urine expressed as osmoles/L and its *specific gravity* (which is the easiest and one of the most useful bedside determinations).

Sp. Gr.[1] . .	1005	1007	1010	1015	1020	1025	1030	1035	1040
Osmoles/L .	0·2	0·3	0·4	0·6	0·8	1·0	1·2	1·4	1·6

[Note that each increment of 0·2 osmoles/L represents an increase of 5 in sp. gr. The concentration of glomerular filtrate=0·3 osmoles/L.]

If the specific gravity of a urine is known the total solute output can be calculated approximately in osmoles (or milliosmoles).

Example. Sp. gr. of urine=1020; this is equivalent to 0·8 osmoles/L.
Daily output of urine=1200 c.c.

∴ Total solute output in 24 hr.=0·8×1·2=0·96 osmoles=960 milliosmoles.

[1] Sp. Gr. 1005, etc., is compared with water which is 1000. If Sp. Gr. of water is taken as 1·000 then Sp. Gr. 1005 is equivalent to 1·005.

Osmotic Work of Tubules. Volume Obligatoire.—The osmotic pressure of urine varies widely; in some conditions, *e.g.* after water drinking it is more dilute than plasma, but generally it is more concentrated. The *average maximal* urinary concentration is 1·4 osmoles/L (plasma=0·3). The tubular epithelium consequently generally performs osmotic work to return water to the blood against the osmotic resistance of the concentrated urine. The urine cannot, however, be concentrated beyond the point at which its osmotic pressure just balances the absorbing power of the renal cells (*i.e.* an osmotic pressure corresponding to a concentration of about 1·4 osmoles/L). The *solutes of the urine therefore demand a certain minimum volume of water for their excretion* (*volume obligatoire*).

If the amount of solute is known, the *minimal* urinary volume in which the solute can be excreted can be calculated. Assume as in the example given above that 960 milliosmoles are to be excreted. At *maximal* urinary concentration, 1400 milliosmoles of solute "demand" 1000 c.c. of urine for their excretion (*i.e.* to give a concentration of 1·4 osmoles/L)

$$\therefore \text{ 960 milliosmoles "demand" } \frac{1000}{1400}\times 960 = 680 \textit{ c.c. of urine.}$$

The Table below shows the volume obligatoire on different diets (yielding different amounts of solutes to be excreted) when urine of different specific gravities is passed.

Diet	Solutes to be excreted in 24 hours (in milliosmoles)	Volume Obligatoire in c.c.		
		Sp. Gr. 1015 0·6 Osmoles/L	Sp. Gr. 1020 0·8 Osmoles/L	Sp. Gr. 1035 1·4 Osmoles/L (maximal concentration)
Usual mixed food intake . .	1200	2000	1500	850
Fasting . .	800	1300	1000	550
Fasting and 100 g. of glucose . .	400	700	500	300
Fasting or a minimal protein intake plus enough glucose to meet full caloric requirements .	200	350	250	150

When plenty of water is available for excretion a larger volume of more dilute urine is passed, *i.e.* a urine of lower osmolarity (lower concentration)

involving less osmotic work in absorbing water in the tubules. When water is scarce the urine is excreted with maximal osmolarity with corresponding maximal osmotic work in the tubules. An early sign of renal failure is impairment of water reabsorption: on a standard water intake a fixed amount of solutes is excreted at lower osmolarity (*i.e.* lower sp. gr. urine) and in a larger volume of urine than is the case with normal people.

Reabsorption of water is only dependent in part on the antidiuretic hormone. Thus in the complete absence of this hormone urine is passed at the rate of 30–40 L per day, *i.e.* out of 170 L of glomerular filtrate 130 L of water (or more) are still reabsorbed (p. 51).

Back-diffusion. Forward-diffusion.—As the osmolarity of the urine rises the concentration of many of the individual constituents in the urine greatly exceeds that in the plasma; there is a tendency for the constituents to diffuse back (*passively*) into the blood; the tubular epithelium must do work to limit this back-diffusion. There is thus a limit to the concentration that may be attained by any *individual* constituent in the urine; for this reason the urea concentration in the urine rarely exceeds 4% (100–200 times as great as the plasma urea concentration).

Generally the concentration of Na^+ and Cl' in the urine is less than in plasma; in *dilute* urine it is very much less. The tubular epithelium must then do work to prevent forward diffusion of Na^+ and Cl' from the plasma into the lumen of the tubule.

Tubular Excretion.—The tubular epithelium can also *actively* transfer substances from the blood into the lumen of the tubules and so eliminate them from the body. This process is called *tubular excretion* (or by an old convention, *tubular secretion*). In mammals none of the *normal* plasma constituents are so excreted (with the doubtful exception of endogenous creatinine). Certain artificially introduced substances are, however, eliminated in this way. Injected (exogenous) creatinine and dyestuffs like diodrast (diodone) and other organic iodine compounds are among the most important substances so treated (p. 37).

Evidence for Mechanism of Renal Secretion.—The evidence for the account of renal secretion given above will now be briefly summarized.

Evidence for Composition of Glomerular Filtrate.—(i) In *amphibia* examination of the fluid in Bowman's capsule collected by *direct puncture* shows that it is identical in composition with the plasma in respect of: crystalloid osmotic pressure, electrical conductivity, alkalinity, and concentration of urea, glucose, inorganic phosphate, creatinine, uric acid, and chloride; the volume of capsular fluid (directly measured) is well above the volume of urine excreted in the same period of time. The capsular fluid in amphibia is thus proved to be a simple filtrate, as assumed on p. 27.

(ii) Similar experiments in *mammals* are far more difficult to perform as the Malpighian corpuscles are placed deep to the renal surface and are covered by loops of proximal tubule. A few successful experiments have been performed in rats and guinea-pigs; the capsular fluid collected by direct puncture was found to be protein (and colloid) free but was otherwise identical in composition with the plasma. The rate of glomerular filtration (determined directly) was 0·38–1·12 c.mm. per glomerulus per hour. It is interesting to note that the human glomerular filtration rate of 120 c.c. per minute is equivalent to a flow of 4·3 c.mm. per glomerulus per hour.

Evidence for Tubular Reabsorption.—(i) In *amphibia* the course of tubular reabsorption can be followed by collecting fluid from different points along the course of the renal tubule. In this way it can be shown that: (*a*) *glucose* is absorbed in the proximal tubule; (*b*) *chloride* is absorbed in the distal tubule; (*c*) *acidification* takes place in a short length of the distal tubule; (*d*) after poisoning with *phloridzin*, glucose reabsorption ceases and the glucose concentration in the proximal tubule rises, proving that *water* is reabsorbed there.

(ii) In *mammals* it is comparatively easy to puncture the proximal tubule,

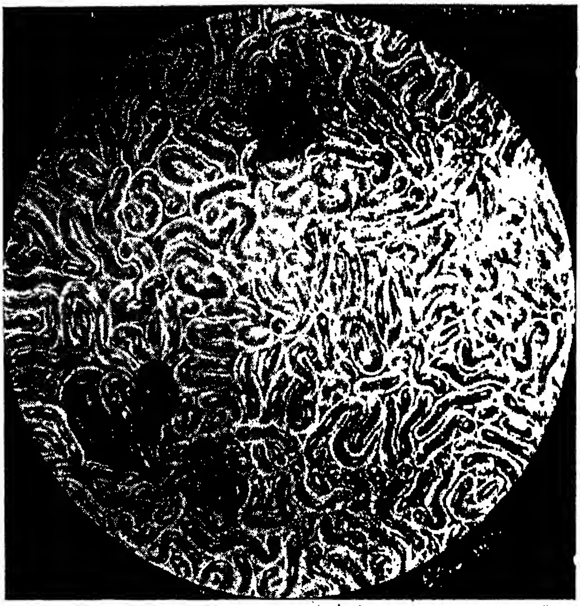


FIG. 10.—Direct Inspection of Living Mammalian Nephron.

Surface view of guinea-pig kidney. The rounded interruption of the tubule pattern at 12 o'clock is a glomerulus. Indian ink has been injected into a single tubule segment and has filled 3 coils of a proximal convolution shown at 7 o'clock. Other tubules visible are also probably proximal tubules. (Walker and Oliver, *Amer. J. Physiol.*, 1941, *134*.)

commonly in its middle third (Fig. 10); very occasionally the distal tubule is struck. Following such an experiment, by an unbelievably skilful technique the entire nephron is dissected out, the exact point of puncture determined and measured off in relation to the entire nephron. The results are illustrated in Figs. 11, 12).

(*a*) *Glucose* is reabsorbed in the proximal tubule; the absolute amount absorbed increases (up to a maximal level) as the plasma glucose concentration is raised. *Phloridzin* paralyses glucose absorption (Fig. 11).

(*b*) *Water.*—Following phloridzin poisoning the glucose concentration half-way along the proximal tubule rises to threefold (which is about as far as the process can be traced) (Fig. 11); probably about 80% of the filtered fluid is reabsorbed in the proximal tubule.

(*c*) *Creatinine.*—Creatinine is neither absorbed nor secreted in the proximal tubule; its concentration rises to the same extent as does that of glucose after treatment with phloridzin (Fig. 11). The rise in the creatinine concentration in the tubular fluid is thus a measure of water absorption.

(*d*) *Chloride.* — The chloride concentration in the first half of the proximal tubule normally rises only by $\times 1{\cdot}5$ (and not $\times 3$, like creatinine). This proves that *partial chloride absorption* takes place here. In the rat the *urinary* chloride concentration may be lower than that of plasma; this proves that the more distal parts of the nephron are mainly responsible for chloride absorption (Fig. 12).

(*e*) *Reaction.* — The reaction becomes more acid owing to reabsorption of $NaHCO_3$; further acidification occurs in the *distal* tubule (pp. 94, 96).

(*f*) *Osmotic Pressure.* — The osmotic pressure remains *unchanged* throughout the proximal tubule; this means, of course, that the fluid absorbed has the same osmotic pressure as the glomerular filtrate. As both chloride and creatinine concentration rise in the proximal tubule and the Na^+ concentration is unaltered, the fluid which is absorbed, though isosmotic with the glomerular filtrate, is of different composition (*i.e.* it contains less chloride and more Na^+ probably as $NaHCO_3$), thus acidifying the fluid. As the osmotic pressure of the urine may be substantially higher than that of the fluid in the proximal tubule, further absorption of water must occur in the *distal* tubule (Fig. 12).

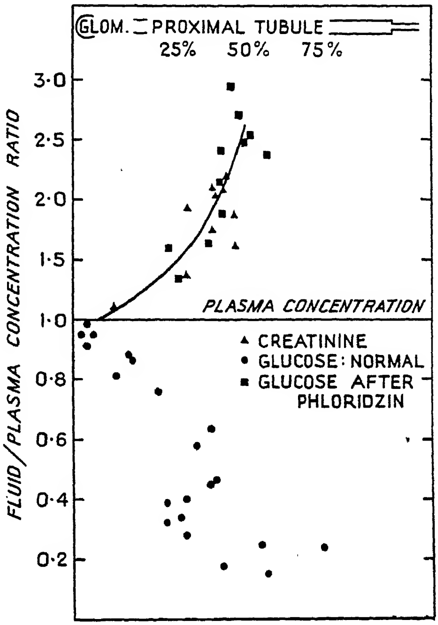


Fig. 11.—Absorption of Glucose and Water in the Renal Tubules. (Redrawn from Walker *et al.*, *Amer. J. Physiol.*, 1941, *134*, 587.)

Experiments on guinea-pigs and rats.
Ordinate: concentration ratios between the fluid collected (by direct puncture) from Bowman's capsule or the proximal tubule, and plasma. 1·0 means that the fluid and the plasma have the same concentration.
Circles; glucose in normal animal.
Black Squares; glucose after poisoning with phloridzin.
Black triangles; creatinine; the results are the same before and after phloridzin.
Note that normally glucose is almost completely absorbed in the proximal tubule.
After phloridzin, glucose and creatinine are concentrated equally owing to absorption of water in the proximal tubule.
25%, 50%, 75% represent points one-quarter, half, and three-quarters the distance along the proximal tubule.

Evidence for Tubular Excretion.—(i) As explained on p. 37, the high plasma clearance values obtained in man for diodrast (700 c.c.) prove that it is largely excreted by the renal tubules.

(ii) In the case of the creatinine normally present in the plasma, the inadequate data available suggest that it is eliminated wholly by glomerular

filtration. If the blood creatinine is raised *in man* by intravenous injection of (exogenous) creatinine, the resulting creatinine clearance value may

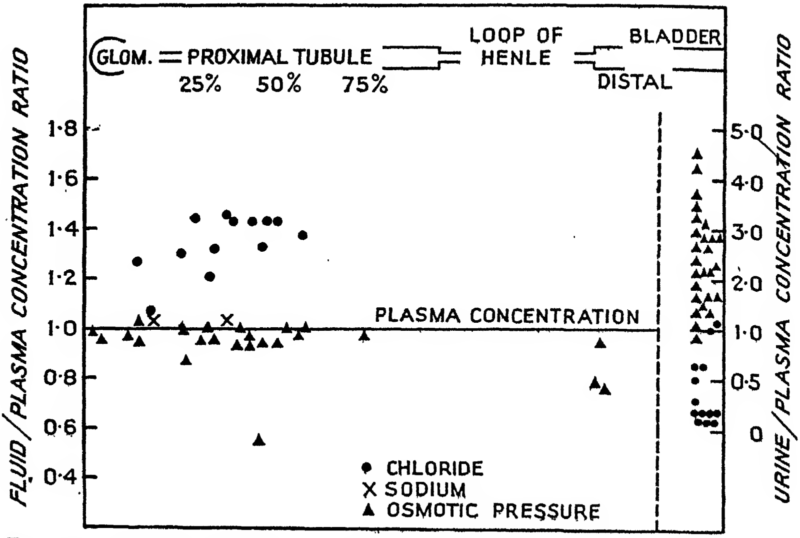


Fig. 12.—Changes in Sodium and Chloride Concentrations and in Total Osmotic Pressure as Glomerular Filtrate passes along Renal Tubule. (Walker *et al.*, *Amer. J. Physiol.*, 1941, *134.*)

Circles = Chloride; *Crosses* = Sodium; *Triangles* = Osmotic pressure.
Vertical axis: left-hand column, F/P = concentration ratio of tubular fluid/plasma. Right-hand coulmn, : U/P = concentration ratio urine/plasma. 25%, 50%, 75% represent points one quarter half-way and three-quarters the way along the proximal tubule.
No change in osmotic pressure or in Na^+ takes place in proximal tubule in spite of absorption of water (Fig. 11); there is an increase in Cl′ concentration (indicating absorption of isosmotic but Cl-poor fluid, containing perhaps excess of $NaHCO_3$). The further absorption of chloride (shown by the fall in Cl′ concentration in spite of absorption of water) and rise of osmotic pressure are due to the absorptive activity of the distal tubule (cf. values for the ultimate urine).

be 170 c.c.; as it then exceeds the inulin clearance value (of 120 c.c.), it is in part excreted by the tubules.

It may be re-emphasized, however, that no other *normal* plasma constituent is excreted by the tubules.

USE OF CLEARANCE VALUES AND OTHER SPECIAL METHODS IN STUDY OF RENAL ACTIVITY

Clearance Value.—The clearance value (C) of a plasma constituent is the volume (in c.c.) of plasma which contains the amount of the constituent which is excreted in the urine in one minute. Consider the clearance value for urea (urea clearance): plasma concentration of urea (P) is 30 mg-%; the amount of urea excreted in the urine in one minute (U_m) is 20 mg.

$$\text{Clearance value (in c.c.)} = \frac{U_m \times 100}{P} = \frac{20 \times 100}{30} = 66 \text{ c.c.}$$

In other words 66 c.c. of plasma contain the amount of urea which is excreted

in the urine in one minute. The term clearance value is somewhat misleading because the *plasma is not cleared of urea.* Of 700 c.c. of plasma which flow through the kidney in one minute, the volume filtered out through the glomeruli is only 120 c.c., containing 36 mg. of urea of which only 20 mg. escape in the urine. The clearance value is thus a so-called " virtual volume "; it is the result of an arithmetical calculation. For all that, the determination of the clearance value for certain substances provides a measure of the volume of the glomerular filtrate and of renal efficiency.

Inulin Clearance. Glomerular Filtration Rate.—The soluble polysaccharide inulin is filtered out from the glomeruli in the same concentration as in plasma; in the tubules the inulin is neither reabsorbed nor excreted. If these statements are true then the inulin clearance value is equal to the volume of the glomerular filtrate.

Suppose the plasma inulin concentration (P) is 100 mg-%, and the inulin excretion in the urine per minute (U_m) is 120 mg. Then

$$\text{Clearance Value} = \frac{U_m \times 100}{P} = \frac{120 \times 100}{100} = 120 \text{ c.c.}$$

i.e. 120 c.c. of plasma have been cleared of inulin.

As inulin is assumed to be neither reabsorbed nor excreted in the tubules, the only way in which 120 c.c. of plasma could have been cleared of inulin is by the filtration of 120 c.c. of (protein-free) plasma through the glomeruli. In other words, the inulin clearance value (120 c.c.) is, as stated above, the glomerular filtrate volume (Fig. 13).

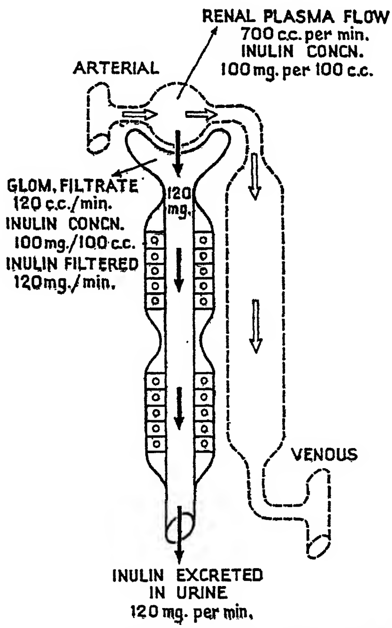


FIG. 13.—Diagram of Inulin Clearance as Measure of Volume of Glomerular Filtrate (Slome).

EVIDENCE THAT INULIN IS ELIMINATED SOLELY BY GLOMERULAR FILTRATION.—Inulin is a soluble polysaccharide with a molecular weight of 5000. If *carefully and thoroughly purified* it can be injected intravenously without ill effect.

(i) In the frog direct examination shows that the inulin concentration in the glomerular filtrate is identical with that in plasma; inulin is *not* excreted by the kidneys of aglomerular fishes.

(ii) In man, *alterations in the plasma inulin concentration (under conditions of constant glomerular filtration rate) do not affect the inulin clearance value.* Thus, if the plasma inulin concentration (P) is doubled, twice the amount of inulin is filtered out from the glomeruli, and twice the amount is passed out in the urine (U_m); as P and U_m rise to the same extent the clearance value is unaltered.

(iii) The clearance values for all substances eliminated solely by glomerular filtration should be the same under constant conditions. In the *dog* the clearance values for inulin, creatinine, and ferrocyanide are identical, proving that they are treated in the same way. In view of the highly selective character of both tubular reabsorption and excretion, it would be most improbable that three such utterly diverse compounds would have the same value, unless in all three cases they were excreted solely by glomerular filtration.

(iv) *Phloridzin* acts directly on the renal tubule, paralysing its power of reabsorbing glucose back into the blood. After injection of phloridzin, the glucose clearance equals the simultaneously determined inulin clearance in the dog, indicating that both substances are treated in the same way (*i.e.* both are purely filtered). *In man* (after poisoning with phloridzin) the glucose/inulin clearance ratio is 0·9; it is not unity because a full paralysing dose of phloridzin cannot be safely given, and so some glucose is still reabsorbed. The value is sufficiently near unity however to indicate that both substances are eliminated under these conditions purely by filtration (Fig. 14). In the *phloridzinised* dog, the clearance values for inulin, creatinine, glucose, xylose, sucrose, and ferrocyanide are identical, *i.e.* they are all exclusively filtered.

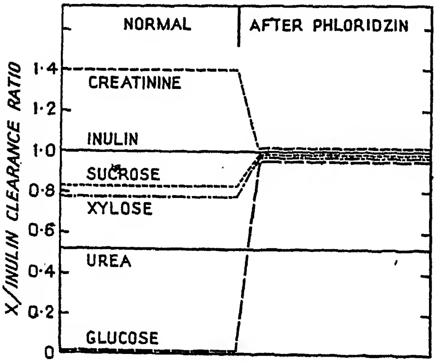


FIG. 14.—Clearance Rates of Various Substances (X) (in *Man*) compared with that of Inulin under Normal Condition (left) and after treatment with Phloridzin (right). (Homer Smith, *Physiology of the Kidney*, 1937).

Where the ratio X/inulin=1, the substance is eliminated by filtration only; if the ratio is over 1, it is also excreted by the tubules; if the ratio is under 1 then reabsorption takes place in the tubules. After phloridzin the ratio in all cases (except urea) is approximately unity.

To *determine glomerular filtration volume in man* a large initial dose of inulin is injected which is followed by a constant inulin infusion at a rate which just compensates for the loss in the urine. A constant plasma level of inulin is maintained. Under these conditions the concentration in the systemic venous blood is for all practical purposes the same as in the arterial blood. The inulin concentration in the plasma of venous blood (P) and the simultaneous inulin excretion per minute in the urine (U_m) are determined. Then:

$$\text{Glomerular Filtrate} = \frac{U_m \times 100 \text{ c.c.}}{P}$$

Measurement of Renal Blood Flow. Diodrast Clearance.—The renal blood flow in animals can be easily measured directly by various means. It can also be determined indirectly by methods based on the Fick principle (as used for determining the pulmonary blood flow and consequently the cardiac output (p. 278)). Let P and V represent respectively the concentration in mg-% of a constituent (A) in the plasma of renal arterial

and renal venous blood ; let U_m be the output in mg. of A in the urine per minute. Thus,

100 c.c. of arterial plasma arrive at the kidney with P, leave with V, and lose P—V ;

P—V is eliminated by 100 c.c. of renal plasma flow.

$$\therefore U_m \text{ is eliminated by } \frac{100}{P-V} \times U_m \text{ c.c.} = \text{renal plasma flow per minute.}$$

The term *extraction* rate is applied to the expression (P—V)/P, and represents the fraction of the arterial plasma concentration (of A) which is excreted by the kidney (cf. coefficient of utilization of oxygen (p. 413)).

This method can be employed in animals using any conveniently determinable plasma constituent (A) ; it involves of course puncture of an artery and of the renal vein. In man, as renal vein puncture is impractical, this method can only be used when the *degree of extraction is constant and is known.* If the extraction from the plasma is *complete*, none of the substance A escapes into the renal venous plasma and V is 0; P—V=P—0=P; the formula for renal plasma flow $\frac{100}{P-V} \times U_m$ then becomes simply $\frac{100}{P} \times U_m$; *i.e.* only arterial blood and urine need be analysed. (As will be mentioned below, under suitable experimental conditions systemic *venous* blood can be used instead of arterial.) From a knowledge of the relative volume of plasma and corpuscles) the *total renal blood flow* can be readily calculated.

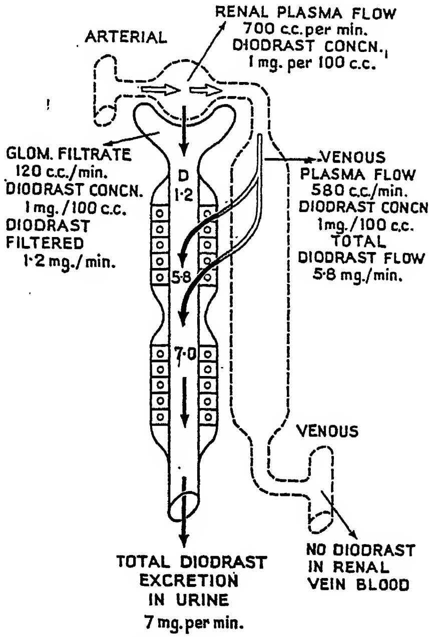


Fig. 15.—Diagram of Diodrast Clearance as Measure of Volume of Renal Plasma Flow. (Slome.)

There is evidence that when certain dyes which are organic iodine compounds, *e.g. diodrast* (diodone), are injected into the blood to produce plasma concentrations not exceeding 5 mg. of contained iodine (or perhaps even 15 mg.) per 100 c.c., they are completely, or almost completely, extracted from the blood during each passage through the kidney. Let us assume a plasma diodrast concentration of 1 mg–% and a renal plasma flow of 700 c.c. per minute. 120 c.c. of glomerular filtrate are formed containing 1·2 mg. of diodrast ; the remaining 580 c.c. of plasma reach the tubules ; their contained 5·8 mg. of diodrast are *actively excreted* by the tubular epithelium from the blood into the lumen of the tubules. In this way *all* the 7 mg. of diodrast

originally present in the 700 c.c. of renal plasma flow are cleared from the blood (cf. Fig. 15). If these statements are true, the diodrast clearance value is a measure of renal plasma flow.[1]

The practical procedure in outline is as follows: after an initial large injection of diodrast intravenously, its administration is kept up at a rate which just compensates for the loss in the urine; under these conditions the systemic *venous* concentration can be used as equivalent for all practical purposes to the arterial concentration, as the concentration in the blood is maintained at a steady level. The diodrast excretion in the urine is simultaneously determined. P (plasma diodrast concentration) and U_m (urinary excretion of diodrast per minute) in the formula $\frac{U_m \times 100}{P}$ being known, the renal plasma flow is easily calculated.

This procedure in man gives a renal plasma flow of 450–700 c.c. and a total renal blood flow of 900–1300 c.c. in relation to a standard surface area of 1·73 sq. metres.

Simultaneous Determination of Inulin and Diodrast Clearance.[2]—To clarify the foregoing discussion data are given below from an experiment in man in which the inulin and diodrast clearances were simultaneously determined (using the formula $\frac{U_m \times 100}{P}$). The results incidentally show that these clearances are unaffected by changes in the volume of urine flow. [The reader is advised to work out the calculations and check the clearance values set out in columns (7) and (8).]

Time after beginning of experiment (minutes).	Urine Volume c.c. per minute.	Inulin		Diodrast Iodine		Clearance Value in c.c.	
		Plasma mg-%.	Urine mg-%.	Plasma mg-%.	Urine mg-%.	Inulin.	Diodrast.
37	8·5	123	1584	0·64	49·6	109	658
47	5·7	122	2235	0·60	67·4	105	641
56	3·6	123	3530	0·58	98·9	103	614
84	2·4	130	5490	0·60	159·0	101	636

Note: To calculate the clearance values, the U_m for inulin and diodrast must first be worked out from the data in columns (2) and (4) and columns (2) and (6) respectively; then apply the formula $U_m \times 100/P$.

Limitations of Diodrast Method of Measuring Renal Plasma Flow.—(i) In *dogs*, 10% of the arterial diodrast escapes into the renal venous blood; as the diodrast clearance is only 90% complete, the use of the formula $\frac{U_m \times 100}{P}$ gives a value for renal plasma flow which is 10% too low. Determination of the diodrast content of the renal venous blood in *man* gives similar results. In so far as the diodrast clearance method in man is inaccurate it gives an underestimate of renal blood flow.

[1] Para-amino-hippurate may be used instead of diodrast and gives similar results.
[2] For simplified clinical procedures, see Olbrich *et al.*, *Lancet*, 1950, ii, 565.

(ii) The diodrast method measures the blood flow only to *normally functioning tubules*. If the blood is passing to non-functioning tubules they do not clear the diodrast and that fraction of the blood flow is not revealed. If the blood were shunted through the medulla—thus mainly by-passing the cortex—it would not reach the epithelium of the *convoluted* tubules in the cortex which are responsible for secreting diodrast from the blood into the lumen of the tubule; here again a fictitiously low value for blood flow would be obtained. This limitation of the diodrast technique is well shown in animal experiments on hæmorrhage where the directly measured blood flow may be nearly twice as great as that calculated from diodrast clearance values. The same error may occur in studies of patients with hæmorrhage or shock.

As will be explained on p. 40, it is possible to assess the functional state of the tubular epithelium by determining the *maximal* powers of the tubules to excrete diodrast (so-called *tubular maximum* or T_m). If the T_m is below normal, the tubules are probably damaged and the diodrast clearance probably gives a correspondingly misleadingly low value for renal blood flow. If the T_m is normal the diodrast clearance is probably a fairly reliable measure of renal blood flow.

Urea Clearance.[1]—The value for urea clearance was calculated on p. 35. The result given there, *i.e.* 66 c.c. has the following meaning. If the glomerular filtrate is 120 c.c. per minute and the plasma urea level is 30 mg-%, then 36 mg. of urea are filtered out per minute through the glomeruli into Bowman's capsule. If 20 mg. of urea are excreted in the urine per minute, then 36—20=16 mg. of urea are reabsorbed into the blood per minute.

Van Slyke's Urea Clearance Test.—According to Van Slyke normal values for urea clearance are partly dependent on the *rate of urine flow*.

(i) If the rate *exceeds* 2 *c.c. per minute*, the usual clearance formula $\frac{U_m}{P}\times 100$ represents the findings. He calls the result the *maximal* urea clearance (C_m) and gives the normal range as 60–95 c.c. (mean 75 c.c.= 100%).

(ii) When the flow of urine *is less than* 2 *c.c. per minute* he uses the formula

$$\frac{\text{Urea \% in urine} \times \sqrt{\text{urinary volume per minute}}}{\text{plasma urea concentration}} = \frac{U \times \sqrt{V}}{P}.$$

The result is the *standard* clearance (C_s) and the normal range is 40–65 c.c. (mean 54 c.c.=100%).

Whatever may be the physiological interpretation of these clearance values, it is claimed by Van Slyke that urea clearance (*expressed as a percentage of the mean normal values*) is a reliable, if empirical, index of the functional state of the kidneys.

Thus, in a series of cases of kidney disease, when the C_s fell to 50% there was no change in blood urea; with C_s between 20 and 40%, in more than half the cases the blood urea was unaltered; only when the C_s was below 20% did nearly all the cases show an elevated blood urea. Similar results were obtained from a study of the blood creatinine; only when the C_s fell below 20% did the blood creatinine exceed the upper normal limit of 2 mg. per

[1] Peters and Van Slyke, *Quantitative Clinical Chemistry, Interpretation, 1*, i, 2nd edn., 1946, 840.

100 c.c. When the urea clearance falls below 5% (of the normal), uræmic symptoms are present; above 10% these symptoms are absent.

The clearance test is carried out as follows: the urine is collected for two one-hour periods; at the end of the first, blood is collected; the volume of urine, its urea concentration, and the plasma urea concentration are determined and the appropriate calculations made.

General Significance of Clearance Values.—(i) When a substance is filtered out in the glomeruli and is neither reabsorbed nor excreted in the tubules, its clearance value equals the volume of glomerular filtrate. This is true of *inulin*.

(ii) When a substance is filtered out in the glomeruli and is reabsorbed in the tubules, its clearance value is less than the volume of the glomerular filtrate, *i.e.* it is less than the inulin clearance value (cf. *urea clearance*, p. 39). The more completely the substance is absorbed, the lower is its normal clearance value.

(iii) When a substance is filtered out in the glomeruli and is also excreted by the tubules, its clearance value exceeds the glomerular filtrate and thus exceeds the inulin clearance value. The more extensively it is excreted in the tubules the higher the clearance value. If a substance is *completely cleared from the plasma, its clearance value equals the renal plasma flow.*

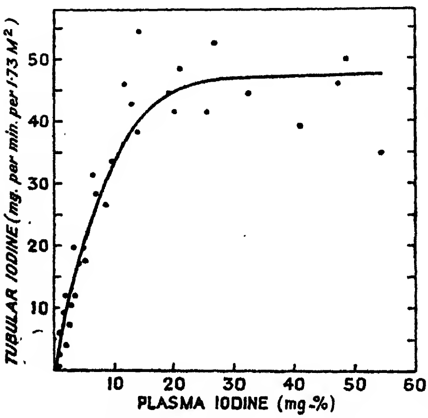


FIG. 16.—Maximal Tubular Excretion of Diodrast Iodine by Kidneys in Man. (Redrawn from White *et al., Proc. Soc. exp. Biol. Med.*, 1943, *43*, 13.)

Diodrast expressed in terms of its iodine content. Ordinate: tubular *excretion* of diodrast (*i.e.* total excretion in urine (U_m) minus amount filtered in glomeruli (F_m)) per standard surface area (1·73 sq. m.). The rate of tubular excretion of diodrast at first rises almost in direct proportion to rise of plasma diodrast level; at plasma level of 15–20 mg. of diodrast I, maximum excretion is reached (T_m). $T_m = 50$ mg./min./1·73 sq. m. surface area.

Determination of Tubular Excretory Power. — Tubular excretion, like tubular reabsorption, depends on the vital activity of the living epithelial cells. Clinically the maximum power of the tubules to excrete foreign substances (*e.g.* diodrast) can be readily determined and is called the T_m or *tubular excretory maximum*). Thus "diodrast T_m" means the maximal power of the tubules to excrete diodrast. This is determined by raising the plasma diodrast concentration well *above the level at which complete extraction is possible*, *i.e.* above 20 mg–% (of iodine); the urinary diodrast excretion is determined (U_m); the amount of diodrast filtered (F_m) in the glomeruli is calculated from the plasma diodrast concentration and the volume of glomerular filtrate (=inulin clearance value) and is deducted; then $T_m = U_m - F_m$. Typical results are shown in Fig. 16. In man the maximal diodrast output is about 50 mg. (of iodine) per minute. The T_m value is related to the *number of functioning tubules and their excretory efficiency*; it is therefore regarded as an index of *tubular excretory mass*. It is uncertain to what extent the T_m for one substance is a guide to the T_m for other unrelated substances, as

specific mechanisms may be concerned in the excretion of each substance or group of substances. Similarly, it is uncertain whether the excretory T_m is a guide to the *reabsorptive* powers of the tubules. In general, however, a *marked decrease in T_m represents a decrease in the number or the general efficiency of the renal tubules (or both combined)*.

Notion of T_m Applied to Glucose Reabsorption.—The notion of T_m can be usefully applied to the treatment of glucose by the renal tubules. The following results were obtained in a dog on progressively raising the plasma glucose concentration.

	Plasma glucose mg–%	Glucose filtered in glomeruli mg./minute	Glucose lost in urine mg./minute	Glucose reabsorbed mg./minute
(1)	126	145	0·15	145
(2)	238	267	2·2	265
(3)	437	499	229	270
(4)	1234	1494	1212	282

Initially as the plasma glucose rose above its normal value all the glucose filtered in the glomeruli was reabsorbed (Table, (1)). But there is an upper limit to the *absolute* amount of glucose that can be absorbed in unit time; in this experiment it was 280 mg./minute. When the absolute amount of glucose filtered out exceeds this value, the balance is lost in the urine (*e.g.* Table, (3) and (4)). The glucose *reabsorption* T_m (or T_{mG}) in this dog was thus 280 mg. In *man* the T_{mG} is thought to be 300–350 mg. per minute. The relation of T_{mG} to clinical glycosuria is fully discussed on p. 927.

STRUCTURE OF PITUITARY GLAND [HYPOPHYSIS].[1] RELATION OF NEURAL DIVISION [NEUROHYPOPHYSIS, POSTERIOR LOBE] TO REGULATION OF WATER AND ELECTROLYTE BALANCE

The pituitary body consists of two parts—a *neural* and a *glandular* division—which, from the standpoint of development, structure, and functions, are entirely distinct organs; it is not clear why they have come to be so intimately associated anatomically. The *structure of both divisions* is described below. The functions of the *neural* division will then be discussed in detail as it is intimately related to the control of water and electrolyte balance. The *functions of the anterior lobe* are considered on pp. 930 *et seq.*

Structure of Pituitary Gland (FIGS. 17 and 18).—(1) NEURAL DIVISION (NEUROHYPOPHYSIS).—The neural division includes the *pars nervosa*, the *infundibular stem* (pituitary stalk), and its continuation into the floor of the third ventricle, called the *median eminence*. The structure of all three parts of the neural division is identical; the outstanding feature

[1] For general reviews of the pituitary, see Assoc. Res. nerv. ment. Dis., *Pituitary Gland*, 1938. Van Dyke, *Physiology and Pharmacology of Pituitary Gland*, Chicago, 1936, *1*; 1939, *2*. Pincus and Thimann, *The Hormones*, N.Y., 1948, *1*; 1949, *2*.

is the very *rich supply of non-medullated nerve fibres* (Figs. 17, 23) which arise in the hypothalamus, mainly in the supraoptic nucleus. Numerous modified neuroglia cells (called *pituicytes*) are present (Fig. 19), with branching processes and nuclei staining with cresyl violet; their significance is unknown. There is a rich capillary blood supply. The absence of epithelial or glandular-looking cells is noteworthy in an organ of proved internally secreting functions (cf. p. 48). The neural division develops as an outgrowth from the *floor of the third ventricle* (*i.e.* it is modified nervous tissue).

(2) Glandular Division [Adenohypophysis].—The glandular division

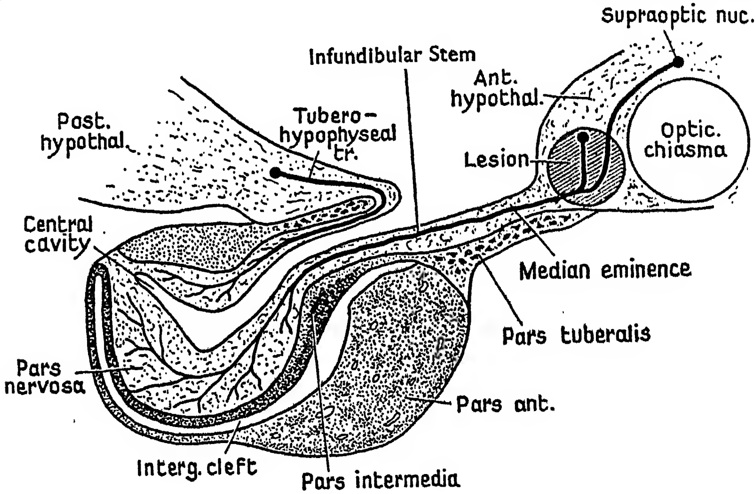


Fig. 17.—Diagram of a mid-sagittal section through the hypothalamus and hypophysis of the cat, showing the two divisions of the hypothalamo-hypophyseal tract, *i.e.* the supraoptico-hypophyseal and the tubero-hypophyseal. The obliquely striped circle indicates the position of a typical lesion designed to produce diabetes insipidus. (Fisher, Ingram, and Ranson, *Diabetes Insipidus*, Edwards Bros. Inc., Ann Arbor, Michigan, 1938.)

Interg. cleft = interglandular cleft; Pars ant. = pars anterior; Post-hypothal., Ant.-hypothal. = posterior and anterior hypothalamus.

develops from the primitive mouth cavity (*stomodœum*); it usually consists of three parts: (i) *pars anterior*, (ii) *pars intermedia* (absent in certain species, and very small in man), and (iii) *pars tuberalis* (rarely found in man).

(i) The *pars anterior* is large, compact, and highly vascular; it consists of cells arranged in nests or columns surrounded by connective tissue. The cells are of two kinds (Plate I):

(a) *Chromophobe*, without affinity for dyes.

(b) *Chromophil*, which stain readily and are subdivided according to the character of their granules into *eosinophil* cells which stain with eosin or acid fuchsin, and *basophil* cells which stain with hæmatoxylin.

The functional significance and mutual relation of these various elements is still obscure; the chromophobe cells may be the undifferentiated non-secreting precursors of the two kinds of chromophil cells. When the latter

PLATE I

SECTION OF PARS ANTERIOR OF NORMAL HUMAN PITUITARY GLAND ($\times$450)

The columns of cells surrounded by the sinusoids are clearly shown. The staining method used (erythrosin-orange G-Toluidene-blue) shows the three types of cell: (*a*) chromophobe, (*b*) chromophil (oxyphil) and (*c*) chromophil (basophil) cells.

Roussy and Oberling. "Contribution à l'Etude des Tumeurs Hypophysaires," *La Presse Medicale.*

[*Facing page* 42

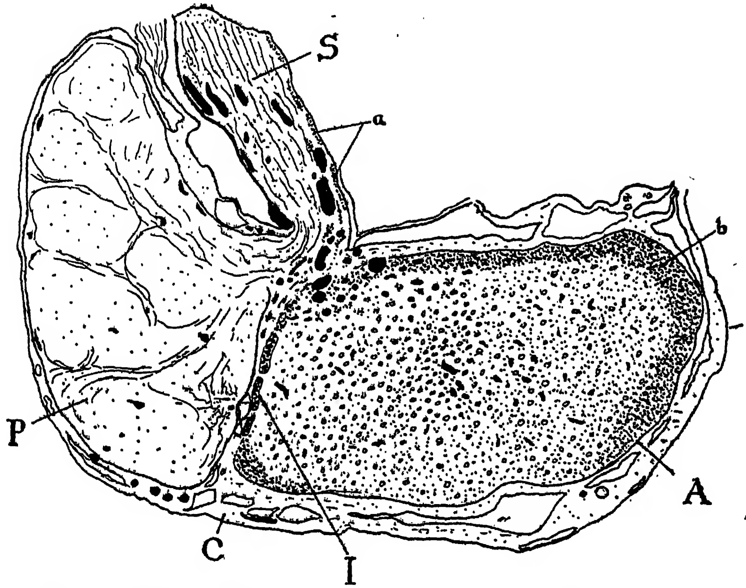


FIG. 18.—Human Pituitary Gland (Sagittal Section). (From a drawing by Prof. J. H. Woodger.)

A=pars anterior. I=pars intermedia. In this specimen it is represented only by three small clumps of cells. *b* (densely stippled)=region of pars anterior which is free from eosinophil cells. The less densely stippled region of the pars anterior contains the eosinophil cells. P=pars nervosa; S=stalk *a*=cells in the stalk resembling those at *b* in the anterior lobe; C=capsule. Blood vessels are indicated in solid black (except the smallest capillaries which are not shown). There is no pars tuberalis present.

FIG. 19.—Histology of Pars Nervosa. (Fisher, Ingram, and Ranson, *Diabetes Insipidus* Edwards Bros. Inc., Ann Arbor, Michigan, 1938.)

A. Transverse section through pars nervosa of normal cat. Cresyl-violet stain. N=pars nervosa, showing nuclei of pituicytes. I (at upper margin of section)=pars intermedia.
B. Normal pars nervosa (higher magnification), showing the pituicytes (*p*). The long processes of the pituicytes and glial fibres can also be seen. Penfield del Rio-Hortega silver carbonate stain.

discharge their secretion they lose their granules and are then supposed to revert to the chromophobe phase.

Useful information can be derived from a study of the effects of innocent tumours (adenoma) arising from the different kinds of cells. Chromophobe cell tumours are associated with signs of hypopituitarism (p. 940); this observation suggests that the chromophobe cells of the tumour form no internal secretion but mechanically destroy the secreting elements in the intact part of the gland. In acromegaly (p. 941) tumours of the eosinophil cells are almost constantly present. The eosinophils probably form growth hormone and prolactin. Changes in the basophil cells occur in Cushing's syndrome (p. 965). [See also pp. 980, 1085.]

Although very few nerve fibres have been traced to the pars anterior there is persuasive evidence that its secretory activity is controlled in part by the central nervous system (p. 931).

(ii) The *pars tuberalis* (see Fig. 17) is difficult to demonstrate in man, though it is conspicuous in the cat where it embraces the infundibular stem and extends to the hypothalamus. When present the pars tuberalis consists of numerous vesicles lined by basophil cells; the vesicles contain coagulated fluid and cellular elements which are probably derived from the degeneration of the surrounding cells. Its functions are unknown.

(iii) The *pars intermedia* is a thin, poorly vascularized tissue of variable structure. In man such little intermedia tissue as is present lies in the pars anterior and consists of densely packed basophil cells. In some species the intermedia is absent. The rôle of the pars intermedia in mammals is unknown.[1]

Anatomically the pituitary is divisible into an *anterior* and a *posterior lobe.* The *anterior lobe* contains the pars anterior of the glandular division; the *posterior lobe* contains the pars nervosa of the neural division.

The Neural Division (Posterior Lobe).[2]—The neural division has two functions:

(i) It secretes the *antidiuretic hormone* which controls the secretion of urine by the kidney and thus regulates the water and electrolyte balance of the body fluids.

(ii) It (probably) secretes the *oxytocic hormone* (*oxytocin*) which stimulates the uterus, especially during parturition. It is believed that oxytocin may be normally related to the onset and progress of parturition. Oxytocin is also probably secreted during lactation and promotes the discharge of milk from the breast.

Extracts of the posterior pituitary contain these two specific hormones. A highly potent preparation of each of these hormones has been isolated which is contaminated to only a small extent by the other. In *physiological* concentrations the antidiuretic hormone and oxytocin act exclusively on the kidney and the uterus (and breast) respectively. In large doses (*i.e.* far larger than are secreted in the body) the antidiuretic hormone produces

[1] In amphibia and certain fishes the pars intermedia forms a hormone which expands the pigment-bearing cells (*melanophores*) in the skin. After hypophysectomy the pigment cells contract and the skin becomes pale; pituitary extracts cause the cells to expand and the skin consequently darkens.

[2] Fisher, Ingram, and Ranson, *Diabetes Insipidus and Neuro-hormonal Control of Water Balance*, Ann Arbor, 1938. Harris, *Physiol. Rev.*, 1948, *28*, 139. Stehle, *Vitamins and Hormones*, 1949, *7*, 383, 390.

striking circulatory effects. *Whole* post-pituitary extracts also act on smooth muscle in many parts of the body. The actions on the circulation and smooth muscle are purely of pharmacological and therapeutic interest and do not represent functions of the neural division.

Actions of Posterior Pituitary Extracts.—The main actions, as explained above, are on (i) the kidney, (ii) the uterus, (iii) the breast, (iv) the circulation, and (v) smooth muscle.

(1) Action on Kidney (Antidiuretic Hormone).—(i) In man on a normal water intake the flow of urine is slightly diminished by post-pituitary extracts (Fig. 20).

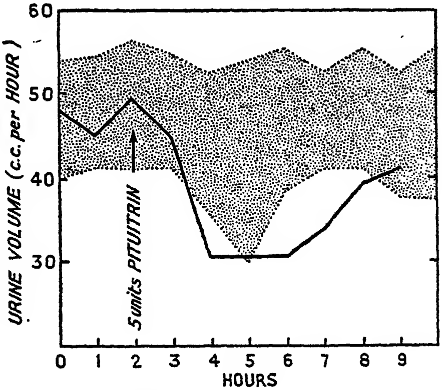


Fig. 20.—Inhibitory Action of Post-pituitary Extract, "Pituitrin," on Normal Urinary Output in Man. (John Marks, *J. Physiol.* 1944.)

Shaded area: range of hourly output of urine (in c.c. per hour) during control period.
At arrow: inject 5 units of posterior pituitary extract.
Note decrease in urinary output below control minima

(ii) The most convincing way of demonstrating the antidiuretic action is as follows: induce a diuresis by drinking a large volume of water; injection of the antidiuretic hormone at the height of the diuresis temporarily cuts down the flow of urine to "resting" levels. Alternatively the antidiuretic hormone is injected while the water is being drunk; the injection then delays for periods which may be as long as 6 or 8 hours the onset of the usual diuresis (Fig. 21).

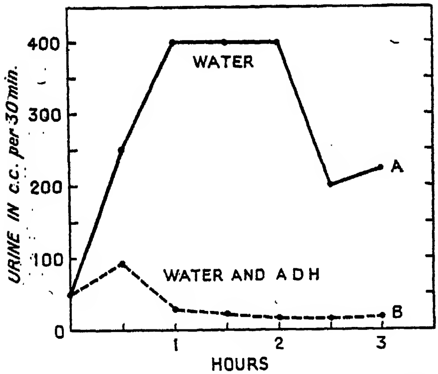


Fig. 21.—Antidiuretic Action of Post-pituitary extracts. (Samson Wright)

A. Ingestion of 1760 c.c. of water on empty stomach. There is a rapid onset of diuresis; about 1900 c.c. are excreted in the urine in the next 3 hours.
B. Ingestion of 1200 c.c. of water combined with an injection (intramuscularly) of 5 units of posterior pituitary extract (antidiuretic hormone=ADH). Diuresis was delayed for 4 hours.

(iii) The antidiuretic hormone probably acts directly on the renal tubules *stimulating the reabsorption of water* and so decreasing the volume of urine formed. The hormone in physiological doses has no effect on the renal circulation or in fact on the circulation anywhere.

(iv) Deficient secretion of antidiuretic hormone is responsible for the disease known as *diabetes insipidus* which is characterized by the passage of a very large volume of very dilute urine (p. 49).

(2) Action on Uterus (Oxytocic Action).—The effects of oxytocin depend on the species and the state of the uterus. In the guinea-pig, oxytocin contracts the isolated virgin uterus suspended in oxygenated Ringer-Locke

solution; this preparation is generally used for the assay of the oxytocic potency of pituitary extracts. In non-pregnant women oxytocin has a feeble stimulating action on the uterus during the first half of the menstrual cycle and a somewhat more marked effect during the second half; oxytocin cannot produce abortion in pregnant women except occasionally late in pregnancy when it is administered combined with quinine and purgatives. But during the *second stage of labour* oxytocin powerfully increases uterine contractions and helps to expel the fœtus; later, it aids the expulsion of the placenta. It also stimulates uterine contractions during the puerperium (Fig. 22). Oxytocic hormone is probably secreted normally by the neural division during labour (p. 1091).

(3) ACTION ON BREAST.—Oxytocin produces contraction of the circularly arranged myoepithelium which surrounds the alveoli of the breast, thus expelling the contained milk into the ducts which are simultaneously kept open by contraction of their longitudinally arranged myoepithelial layer (cf. p. 1094). Oxytocin does not increase the amount of milk formed, but it helps the expulsion of the milk which is already present in the breast.

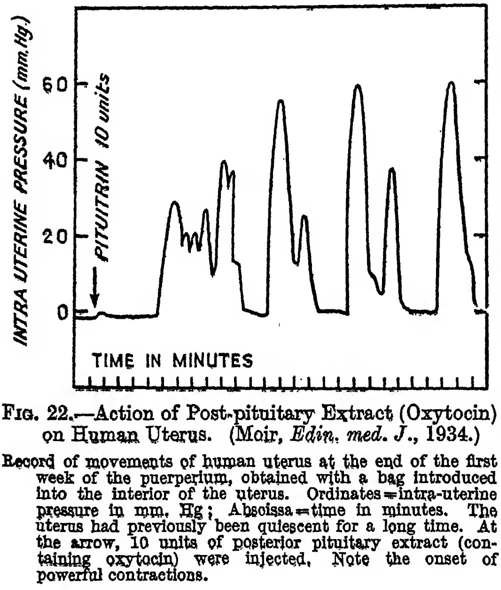


FIG. 22.—Action of Post-pituitary Extract (Oxytocin) on Human Uterus. (Moir, *Edin. med. J.*, 1934.)

Record of movements of human uterus at the end of the first week of the puerperium, obtained with a bag introduced into the interior of the uterus. Ordinates = intra-uterine pressure in mm. Hg; Abscissa = time in minutes. The uterus had previously been quiescent for a long time. At the arrow, 10 units of posterior pituitary extract (containing oxytocin) were injected. Note the onset of powerful contractions.

(4) ACTION ON CIRCULATION.—The effects to be described here and in (5) do *not* represent physiological actions exerted by the gland in the body.

(i) *Heart.*—In animals the rate of the heart is diminished, partly reflexly owing to the rise of blood pressure, and partly by a direct peripheral action; the effect on the *force* is variable. The coronary arteries are constricted. The *human* heart is slowed; the effect on the force is not known.

(ii) *Blood Vessels.*—The *arterioles* in animals are contracted by a direct peripheral action on their walls; after a variable initial change the *blood pressure* rises gradually to a considerable height and declines equally slowly.

In the *human* subject, subcutaneous or intramuscular injection of 5–10 units of post-pituitary extract produces variable effects on the blood pressure. The pressure sometimes rises slightly; much more usually a *fall* is obtained, which may be 20 mm. Hg or more in extent. The most marked circulatory effect in man is a *direct* action, *not* on the arterioles but on the *capillaries*, which contract vigorously. This observation makes the fall of blood pressure particularly difficult to explain. With subcutaneous injections, *local* pallor of the skin results from local active capillary constriction. With intramuscular

injections, more generalized capillary constriction occurs affecting especially the skin of the face (p. 322).

(5) OTHER ORGANS.—Whole post-pituitary extract stimulates smooth muscle widely throughout the body; this action is a direct one and is independent of the integrity of the nerve supply of the tissues concerned. The smooth muscle of the bladder and of the small and large intestine is stimulated. The action on bronchial muscle is still in dispute; probably it is relaxed.

Nervous Control of Neural Division.—As already mentioned (cf. Fig. 17), the neural division receives many nerve fibres (all non-myelinated)

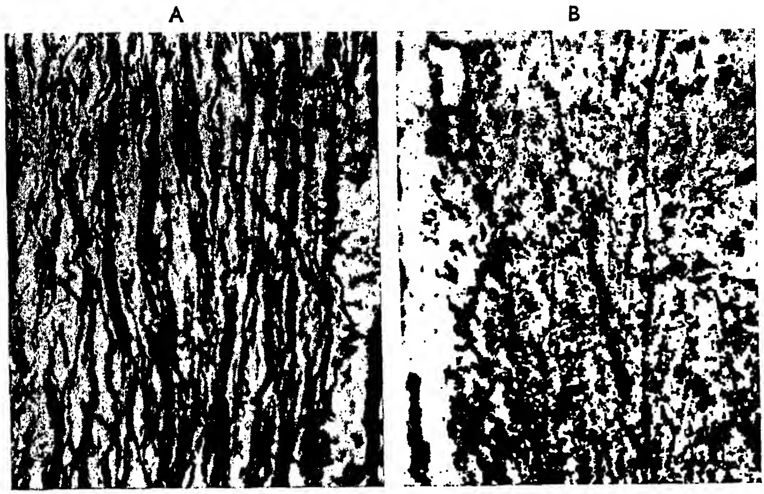


FIG. 23.—Degeneration of Supraoptico-hypophyseal Tract after appropriate Hypothalamic Lesion. (Fisher, Ingram, and Ranson, *Diabetes Insipidus*, Edwards.)

A. Sagittal section through ventral wall of infundibular stem of hypophysis of normal cat stained with pyridine-silver to show the density of the non-medullated fibres constituting the supraoptico-hypophyseal tract.

B. Section of infundibular stem of cat with experimental diabetes insipidus showing degeneration of the fibres of the supraoptico-hypophyseal tract. Only a few abnormal-looking fibres have survived.

from the hypothalamus, chiefly from the supraoptic nucleus and the tuber-cinereum. [Some fibres may also pass into the pars intermedia.] The fibres are well displayed in Fig. 23, A as they run through the infundibular stem; in Fig. 24, A they are seen to form a dense plexus of fibres in the pars nervosa itself. If the nerve tracts are severed by a suitably placed *lesion in the hypothalamus* these fibres undergo degeneration throughout their course: Fig. 23, B shows the disintegration of the fibres as they run through the infundibular stem; in Fig. 24, B practically all the nerve fibres in the pars nervosa have disappeared. The neural division also undergoes other important changes: it atrophies markedly (Fig. 25); the pituicytes disappear; and there is a great increase in cellularity (Fig. 24, B) owing to the development of new tissue, the nature of which is not understood.

The *secretory activity of the neural division is wholly under nervous control.* Following the nerve lesions just mentioned hormone formation and secretion ceases. Extracts of the atrophic denervated gland display no antidiuretic

oxytocic or other activity. The denervation operation is also followed by marked signs of functional deficiency, principally those of *diabetes insipidus* and *failure of parturition* (*dystocia*) (p. 49).

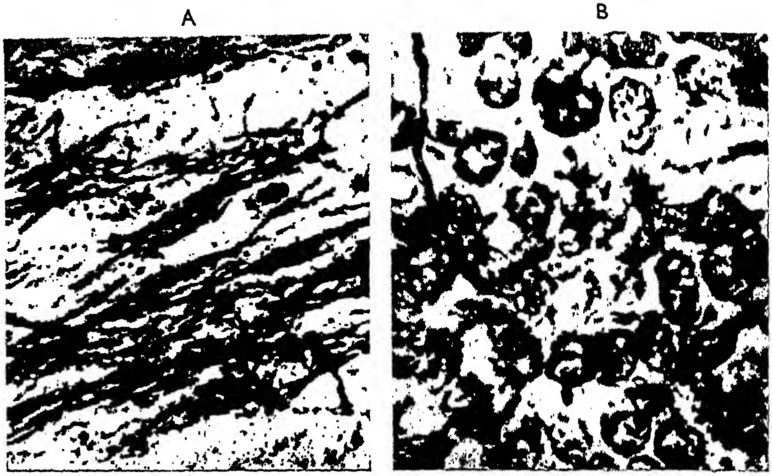


Fig. 24—Degeneration of Pars Nervosa in Experimental Diabetes Insipidus. (Fisher, Ingram, and Ranson, *Diabetes Insipidus*, Edwards.)

A. Transverse section through pars nervosa of monkey (normal) showing dense plexus of non-medullated fibres.
B. Transverse section through atrophic pars nervosa of monkey resulting from hypothalamic lesion interrupting its nerve supply. (The animal suffered from experimental diabetes insipidus.) Note the marked hypercellularity. All the nerve fibres have disappeared except one in the upper left-hand corner of the field. Pyridine-silver stain.

It seems strange at first sight that a tissue completely devoid of glandular cells should be capable of forming several hormones. But this objection from histology is less disturbing in view of the release of specific chemical transmitters at nerve endings which also contain no glandular tissue. The neural

Fig. 25.—Changes in Pituitary in Experimental Diabetes Insipidus. (Fisher, Ingram, and Ranson, *Diabetes Insipidus*, Edwards.)

Left: Transverse section through norma hypophysis (cat). Cresyl-violet stain. A = pars anterior; I = pars intermedia; N = pars nervosa; C = central cavity of pars nervosa.
Right: Transverse section through hypophysis of cat with experimental diabetes insipidus. Note the great atrophy of the pars nervosa and its deeper staining, the enlargement of the central cavity, and the widening of the interglandular cleft (I.C.).

division may perhaps be regarded as equivalent to a massive nerve ending forming hormones exclusively under nervous influences ; these hormones do not act *locally* (like the usual chemical nerve transmitters) but *enter the blood stream* and produce widespread and manifold effects.

Relation of Neural Division to Diabetes Insipidus.—Diabetes insipidus in man is a disease characterized by an increased urinary output which commonly exceeds 10 litres daily ; a daily secretion of 40 litres has been recorded. The urine is very dilute but contains no abnormal constituents. The condition should not be confused with diabetes mellitus, in which the polyuria is associated with glycosuria.

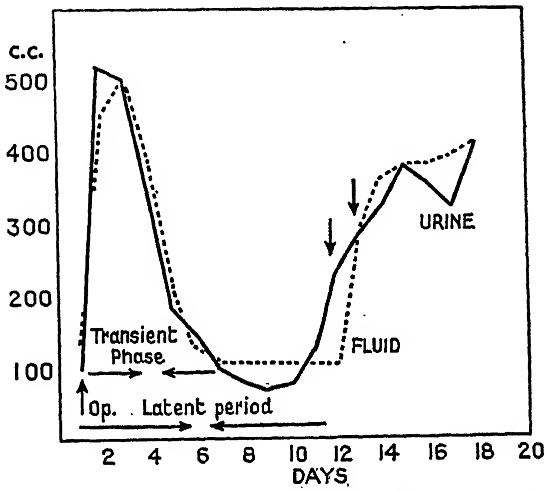


FIG. 26.—Urine Flow and Water Intake in Experimental Diabetes Insipidus resulting from a lesion of the Hypothalamus. (Fisher, Ingram, and Ranson, *Diabetes Insipidus* Edwards Bros. Inc., Ann Arbor, Michigan, 1938.)

Ordinate represents daily urinary output and daily water intake in c.c.
Curves showing urine output (continuous line) and fluid intake (dotted line) in typical experimenta. diabetes insipidus in the cat. Op=time of operation. Note initial transient polyuria, return to normal (together forming the latent period), followed by onset of lasting polyuria (first inverted arrow) and persistent excessive fluid intake or polydipsia (second inverted arrow). The permanent polyuria sets in about one day before the onset of the polydipsia.

EXPERIMENTAL DIABETES INSIPIDUS.—This condition develops in cats and monkeys when lesions are produced in the hypothalamus interrupting bilaterally the pathways from the supraoptic nuclei to the pituitary. Unilateral lesions, or injuries which spare these tracts, do not give rise to polyuria. The severity of the symptoms is proportional to the completeness of destruction of the relevant nerve fibres.

The normal daily urinary secretion in a cat is about 100 c.c. Following an appropriately placed lesion there is (i) an initial transient polyuria (lasting about 5 days), (ii) followed by the return of the urinary flow to its pre-operation level ; the mechanism of these two phases is not understood ; (ii) after 10–14 days a lasting polyuria develops, the so-called *permanent phase* of diabetes insipidus, which is due essentially to lack of antidiuretic hormone (Fig. 26) ; the urinary output may then be 450 c.c. daily. Owing to the polyuria great

thirst develops which leads to ingestion of large quantities of water (polydipsia) to make good the water balance of the body. The experimental state of diabetes insipidus persists with fluctuations for months; had the animals been allowed to survive the condition might have continued indefinitely.

Mechanism of Permanent Polyuria.—The polyuria is due to *lack of antidiuretic hormone, associated with the presence of an intact anterior lobe.* Lack of antidiuretic hormone prevents adequate reabsorption of water in the renal tubules. The evidence is as follows:

(i) It has already been pointed out that the neural division after section of its nerve supply is (*a*) atrophic (Fig. 25), and (*b*) devoid of its usual store of hormones. Like the adrenal medulla, the neural division only forms its hormones or secretes them when its nerve supply is intact.

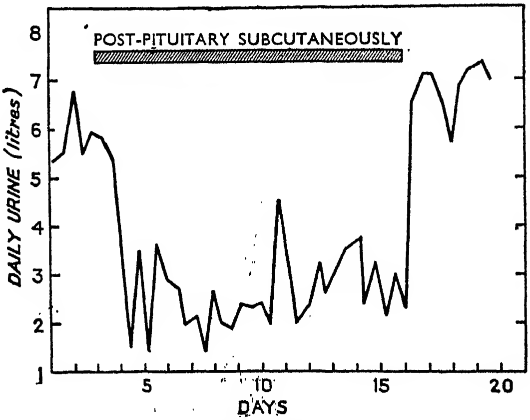


Fig. 27.—Treatment of Clinical Diabetes Insipidus with Posterior Pituitary Extract. (After Christian, from Cushny, *Secretion of Urine*, Longmans, Green & Co., 1926.)

Ordinate: Daily volume of urine in child with diabetes insipidus. Abscissa: Time in days.
During the period indicated "pituitrin" was injected subcutaneously, at first 0·25 c.c. thrice daily, later 0·05 c.c. twice daily. The average volume of urine fell from 6 litres to 2·5 litres daily and rose again when treatment was discontinued.

(ii) *Removal* of the posterior lobe should on this theory regularly produce diabetes insipidus; actually, polyuria commonly occurs but not invariably. The reasons for the exceptions are twofold: (*a*) removal of the pars nervosa still leaves behind histologically identical tissue in the infundibular stem and the median eminence; the persistence of as little as 5% of the tissue of the neural division can maintain the resting urinary output within normal limits. (*b*) The operation may involve coincident damage to the *anterior* lobe or interference with its blood supply.

(iii) *Clinically*, diabetes insipidus occurs with lesions of the posterior lobe, but only when the anterior lobe is intact wholly or partially. In one case of posterior lobe tumour the polyuria diminished as the disease progressively encroached on the anterior lobe.

(iv) *Total hypophysectomy* does not cause diabetes insipidus; on the contrary the diuresis following water drinking becomes markedly delayed, the excess water being excreted in 24 (instead of 3) hours.

(v) The polyuria of both experimental and clinical diabetes insipidus

can be completely controlled by suitable doses of posterior lobe extracts (antidiuretic hormone) (Fig. 27).

Clinical diabetes insipidus is likewise due to lack of antidiuretic hormone resulting from (i) injury or disease of the secreting tissue (*i.e.* the neural division and especially the pars nervosa) or (ii) from degeneration and cessation of activity of this tissue as a result of cutting off its nervous control from the hypothalamus owing to suitably situated lesions in that region.

Rôle of Anterior Pituitary.—The relationship of the anterior pituitary to the kidney is obscure. Following total hypophysectomy (in dogs) glomerular filtration rate, renal plasma flow and maximal diodrast excretion are all markedly reduced and water diuresis is delayed[1] (cf. p. 930). Some conclude that the anterior lobe secretes an unidentified *diuretic* hormone; others suggest the anterior lobe acts more indirectly by its complex effects on general metabolic processes. Injection of pure ACTH (p. 964), however, causes water and salt retention.

Rôle of Thyroid.—The thyroid gland may also be related to the disease. Thyroidectomy reduces the urinary flow in diabetes insipidus, though not to the pre-operative level. If thyroid extract is then administered the urinary flow increases enormously and may persist at a high (though not at the peak) level for a long time after the extract is discontinued. Thyroid extract has a diuretic action in normal subjects; the urinary flow is diminished in myxœdema. All these observations indicate that the thyroid may influence renal secretion, probably not directly by an action on the kidney, but through its influence on metabolic processes in general.

Rôle of Antidiuretic Hormone in Tubular Reabsorption.—The facts set out above show that the antidiuretic hormone is responsible for only a portion and not the whole of the reabsorption of water in the renal tubules. Thus in man about 170 litres of capsular fluid are normally filtered out from the glomeruli in 24 hours, of which 168·5 litres are reabsorbed into the blood and 1·5 litres escape as urine (cf. p. 28). The maximum urinary output in the most severe cases of diabetes insipidus is about 40 litres, indicating that about 130 litres of water are still reabsorbed. Similar results are also obtained in the experimental disease.

Regulation of Secretion of Antidiuretic Hormone (ADH).[2]—The secretion of this hormone is regulated—

(i) By the *emotional* state: exercise, anger, fear, and nocuous afferent stimulation increase the secretion of ADH.

(ii) Mainly by the *crystalloid osmotic pressure* of the plasma. Changes in this osmotic pressure specifically affect sensitive receptors (*osmoreceptors*) lying in the distribution of the internal carotid artery; from these receptors impulses pass up to the hypothalamus and thence to the neural division modifying the secretion of ADH. Increased osmotic pressure (*i.e.* hypertonicity of the plasma) increases, and decreased osmotic pressure (*i.e.* hypotonicity of the plasma) decreases the secretion of ADH.

In the experimental study of the factors controlling the secretion of antidiuretic hormone, its release is demonstrated and the amount is assayed by the effect produced on a *background of diuresis* induced by water drinking. The experimental findings are reviewed in detail below.

[1] Pickford and Watt, *J. Endrocin.*, 1950, *6*, 398.

[2] Verney, *Proc. roy. Soc. B.*, 1947, *135*, 25. O'Connor, *Quart. J. exp. Physiol.*, 1950, *36*, 21.

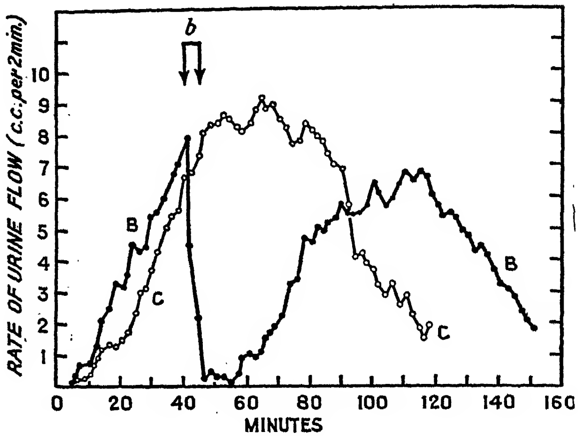


Fig. 28.—Effect of Exercise on Flow of Urine. (Dog). (Verney, *Lancet*, 1946, ii, 740.)

Curve C: At O, administer 250 c.c. of water by mouth. Note the resulting diuresis
Curve B: Repeat water drinking, but during *b* (double arrows) the animal carried out "stationary running" on a platform moving at 6 miles per hour. Note the transient inhibition of the diuresis.

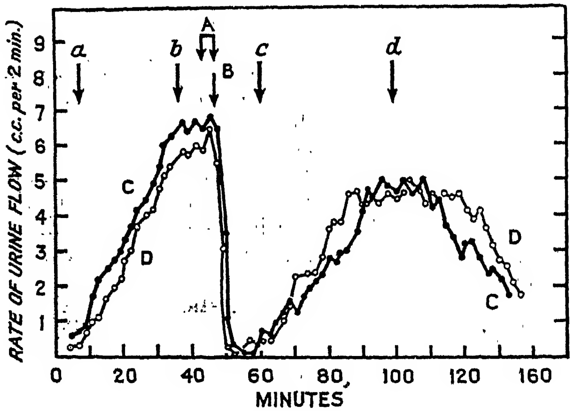


Fig. 29.—Effect of Emotion compared with that of Injection of Antidiuretic Hormone on Flow of Urine. (Dog). (Verney, *Lancet*, 1946, ii, 743.)

Both kidneys and the left adrenal had been denervated and the right adrenal removed.
Curves C and D show the diuretic response to water drinking at O.
At the arrow B on curve C, 5×10^{-4} c.c. of post-pituitary extract (10 units/c.c.) were injected intravenously. Note the antidiuretic effect.
During the double arrows marked A on curve D the animal was angered by noise. Note the remarkable similarity of the resulting antidiuretic effect with that produced by post-pituitary extract.
At the arrows *a*, *b*, *c*, *d*, samples of urine were collected and analysed with the following results:

	a		*b*		*c*		*d*	
	NaCl	Urinary N	NaCl	Urinary N	NaCl	Urinary N	NaCl	Urinary N
Curve C . .	130	430	10	110	80	240	10	130
Curve D . .	160	320	15	150	70	210	5	100

Urinary NaCl and Urinary N are expressed as mg. per 100 c.c. of urine. Note the dilution of the urine at the height of the diuresis (*i.e.* sample *b*) and with the return of the diuresis (sample *d*).

1. **Effect of Emotional States and Exercise.**—If exercise is carried out at the height of water diuresis the flow of urine is temporarily cut down (Fig. 28). This antidiuretic effect is still obtained in animals after severing the sympathetic supply to the kidneys and denervating the adrenals (Fig. 29); it is thus not due to reduced renal blood flow resulting from sympathetic overactivity or adrenaline secretion; in fact, direct measurement shows that no decrease in renal blood flow takes place. The antidiuretic effect is not due directly to the muscular activity, because if the animal is repeatedly exercised the antidiuretic effect dwindles and is finally extinguished. The effect is due to the *emotional concomitant* of the exercise: a similar antidiuretic action is readily produced by fear, anger, or nocuous stimulation. The following observations show that the antidiuretic effect in these circumstances is due to release of ADH:

(i) An exactly similar effect in extent and time course can be produced by injecting an appropriate dose of ADH; this amount is presumably equal to the quantity released by the gland in the experimental procedure (Fig. 29).

(ii) The antidiuretic response produced by emotional disturbance is diminished by removal of the neural division to an extent which is proportional to the amount of neural tissue removed. After removal of the posterior lobe only (in the cat) the response is cut down to 10% or less, the small effect that is still obtained being due to persistence of the infundibular stalk and the median eminence.

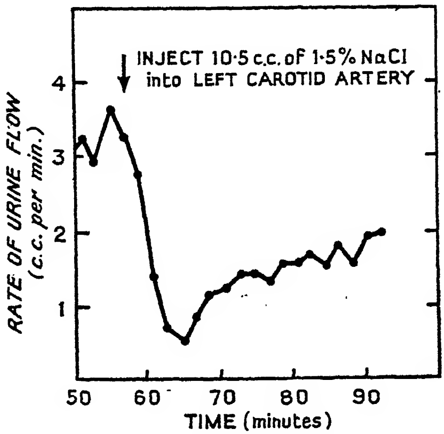


FIG. 30.—Effect of Intra-arterial Injection of Hypertonic NaCl Solution on Flow of Urine. (Dog.) (Verney, *Proc. roy Soc. B.*, 1947, *135*, 59.)

The abscissa shows time in minutes after drinking a test dose of water which induced diuresis. Inject at arrow, 10·5 c.c. of hypertonic saline (0·257M-NaCl, *i.e.* 1·5% NaCl) into the distal end of the carotid artery. Note the immediate inhibition of the diuresis owing to secretion of antidiuretic hormone.

Emotional states or nocuous stimulation presumably act on the neural division via the hypothalamic nuclei (cf. p. 663). The value to the organism of the antidiuretic response is not clear, but it would tend to conserve water in states of stress and danger.

2. **Effect of Changes in Plasma Crystalloid Osmotic Pressure.**—(1) EFFECTS OF RAISED OSMOTIC PRESSURE.—Injection into the carotid artery of a hypertonic NaCl solution produces an immediate antidiuretic response (Fig. 30); using the methods described above, it can be shown to be due to release of ADH. The effective stimulus is the increase in the crystalloid osmotic pressure of the (carotid) arterial blood. Thus:

(i) Solutions of NaCl, Na_2SO_4, or sucrose of equal osmolar concentration produce effects of equal magnitude.

(ii) The amount of ADH released is directly proportional to the degree of hypertonicity induced in the carotid blood.

Verney attributes the results to stimulation of specific *osmoreceptors*: these are as yet undemonstrated receptors (and therefore of unknown structure) which are believed to be specifically sensitive to changes in the crystalloid osmotic pressure. The osmoreceptors set up impulses which reflexly act on the hypothalamus and so on the neural division. The osmoreceptors do not lie in the carotid body (*i.e.* they are not identical with the chemoreceptors (p. 738), but are probably distributed somewhere in the vascular bed of the internal carotid artery. It is a little dangerous to speculate about the mode of stimulation of histologically unidentified receptors, but the following plausible suggestions may be made. It is supposed that the osmoreceptors are surrounded by a membrane which is relatively or absolutely impermeable to Na^+, Cl', SO_4'' or sucrose. Injection of these substances into the blood raises the plasma osmotic pressure; as these substances cannot penetrate into the osmoreceptors a difference of osmotic pressure is established between the plasma and the osmoreceptors. Water is withdrawn from the latter and the resulting local rise of osmotic pressure is the effective stimulus to the sensory nerve endings. A rise of plasma osmotic pressure resulting from excess of a substance to which the membrane is partly or wholly permeable would have less or no effect, because the substance would traverse the membrane and equalize the osmotic pressure on the two sides. It is found experimentally that intracarotid injection of a strong *urea* solution produces no antidiuretic effect; injection of an equiosmotic *glucose* solution releases less ADH than does a NaCl solution. The osmoreceptor membranes are therefore presumed to be freely and quickly permeable to urea but less rapidly and readily permeable to glucose. The osmoreceptors respond steadily (as judged by release of ADH) to a *sustained* increase of osmotic pressure for a period of at least 40 min.; *i.e.* they adapt (accommodate) slowly (cf. p. 550).

The physiological significance of these observations is clear. Normally a *rise* of crystalloid osmotic pressure might be due to (*a*) water deprivation, (*b*) simple salt (Na^+, Cl') excess, (*c*) excess of some other solutes (*e.g.* urea, glucose).

(i) In the case of *water deprivation*, release of ADH cuts down the flow of urine to the obligatory minimum (cf. p. 30), thus conserving water and so tending to preserve the normal plasma osmotic pressure.

(ii) In conditions of *salt excess*, though the NaCl output in the urine increases greatly, the urine flow increases only moderately (p. 62); increased secretion of ADH may be reducing the "tubule diuresis" that the excess NaCl in the urine would otherwise produce. Any water drunk subsequently is mainly initially retained (again possibly by the action of ADH), although it increases the volume of the body fluids; it helps, however, to lower the crystalloid o.p. towards normal.

The osmoreceptors do not respond to intracarotid injection of an *isotonic* salt solution, though it increases the plasma volume and dilutes the plasma proteins. It is interesting that the osmoreceptors respond with maximal sensitivity to the two principal "structural" ions of extracellular fluid, namely Na^+ and Cl'.

(iii) (*a*) An increase of blood *urea* normally leads to increased urea secretion by the kidneys; the elimination of urea is facilitated by an increase in urine volume and is correspondingly hampered by a decrease in urine volume. The non-responsiveness of the osmoreceptors to raised blood urea means that they will not "hamper" urea excretion by decreasing the secretion of urine via the release of ADH.

(*b*) An increase of blood *sugar* such as occurs normally after meals does not call for any urinary adjustment as glucose (unlike Na^+, Cl') is normally metabolized and so got rid of. In diabetes mellitus there is a sustained rise of blood sugar; the kidney responds by excretion of sugar in the urine (glycosuria) which is facilitated by a polyuria. The comparative insensitivity of the osmoreceptors to raised blood sugar means that ADH will not interfere much with the polyuria of diabetes mellitus (p. 921).

(2) Effects of Lowered Osmotic Pressure.—(i) It is reasonable to suppose that a *decrease of crystalloid osmotic pressure, i.e.* dilution of the plasma, would reflexly decrease or arrest secretion of ADH. The enormous diuresis which occurs after water drinking (and which is associated with a detectable decrease in crystalloid osmotic pressure) can be most readily accounted for on these lines; *i.e.* water diuresis is a transient state of diabetes insipidus (p. 58).

(ii) Simple *salt deprivation* leads to dilution of the blood but *not to a diuresis* (p. 65). It seems possible that when the Na^+ and Cl' change develops *very* gradually the osmoreceptors *do* become adapted and no longer respond in their normal manner. When in cases of simple salt lack water is administered (and a *sudden* further osmotic change is produced in the blood) the osmoreceptors do respond but inadequately as judged by the delayed and drawn out water diuresis (Fig. 41).

Clinical Tests of Hypothalamo-Hypophyseal Function.[1] — A secretion of antidiuretic hormone resulting from stimulation of the hypothalamo-hypophyseal mechanism can be produced clinically in the following ways: (i) by smoking one or two cigarettes or injection of nicotine; (ii) by injection of acetylcholine; (iii) by large intravenous infusions of hypertonic saline. If these procedures do not produce an initial falling off in an induced water diuresis or do not temporarily diminish a clinical polyuria there is some abnormality present in the *nervous control* of the post-pituitary or some abnormality in the *post-pituitary* itself. Clinical polyuria of hypothalamo-hypophyseal origin can thus be distinguished from polyurias due to other causes.

Regulation of Oxytocin Secretion.—(i) There is evidence that stimulation of the hypothalamus can elicit a secretion of oxytocin.

(ii) The act of suckling probably reflexly causes a discharge of oxytocin which helps the expulsion of milk from the breast (p. 1094).

(iii) The possible release of oxytocin at term and its rôle in parturition are considered on p. 1091.

(iv) No function has yet been assigned to the oxytocin which is present in the male pituitary in the same concentration as in the female.

EFFECTS OF WATER AND SALT EXCESS AND LACK. REGULATION OF WATER BALANCE AND COMPOSITION OF BODY FLUIDS BY KIDNEY[2]

Effect of Water Drinking: Water Diuresis.—If 1–2 litres of water are drunk, particularly on an empty stomach, absorption takes place rapidly

[1] Chalmers and Lewis, *Clin. Sci.*, 1951, *10*, 127, 137.

[2] Gamble, *Extracellular Fluid*, Harvard, 1949. Marriott, *Brit. med. J.*, 1947, i, 245, 285, 328.

in the extracellular fluid (instead of being, as normally, the same). As the cell membranes behave as though they were impermeable to electrolytes, no net diffusion of these substances out of the cells into the interstitial fluid can occur. Owing to the higher o.p. of the intracellular fluid water is sucked from the interstitial fluid into the cells (Fig. 31).

(iii) Finally the volume of *all* the compartments of body water is increased to the same slight degree, and all the body fluids have an identical but slightly lowered crystalloid o.p. Were there no kidney response the result would be that instead of the 2 L of water that were absorbed being stored exclusively in the plasma, increasing its volume from 3·5 to 5·5 L (*i.e.* by 70%), the water is distributed through the entire 50 L of body water, increasing its volume to 52 L, or by 4%; the plasma volume is likewise increased only by 4%.

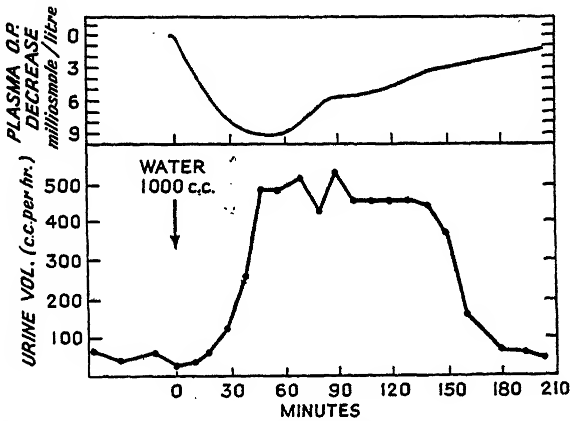


FIG. 32.—Changes in Urine Volume and Plasma *Crystalloid* Osmotic Pressure after Drinking 1000 c.c. of Water in Man. (After Baldes and Smirk, *J. Physiol.*, 1934, *82*.)

Urine volume in c.c. per hour. Change in plasma crystalloid osmotic pressure in milliosmoles per litre. The maximal fall in plasma crystalloid osmotic pressure is about 10 milliosmoles per L. The dilution of the plasma precedes by about 15 minutes the onset of diuresis.

Similarly, instead of the crystalloid o.p. of the plasma being reduced by 70% it falls only by 4%.

(2) RENAL CHANGES.—The kidneys come into action after a latent period of 15–30 minutes (Fig. 32); the flow of urine rises (from the "resting" value of 50 c.c. per hour) to its peak during the second hour when a maximum excretory rate up to 1300 c.c. per hour may be attained; the diuresis declines and is usually over in 3 hours by which time the excess urinary output has about equalled the excess fluid intake. Even 5 litres of water drunk during 2 hours may be eliminated in 4–5 hours.

As the volume of urine increases, its specific gravity falls, *e.g.* to 1001; there may be a slight *total* increased excretion of NaCl and urea during the diuresis, though the *percentage* concentration is very low. This initial washing out of solids is compensated for by a lessened rate of excretion after the diuresis is over. The point to emphasize, however, is that the kidney responds selectively by an enormous increase in water output with little associated loss of solids. The changes in renal dynamics are illustrated in

Fig. 33. There is no increase in renal blood flow ; there is no increase in the volume of glomerular filtrate except when the volume of urine exceeds 900 c.c. per hour ; it is quite clear therefore that the diuresis is due to *decreased reabsorption of water by the renal tubules.* A urinary output of 900 c.c. per hour (=15 c.c. per minute) means that of every 120 c.c. of glomerular filtrate, 105 c.c. are reabsorbed and 15 c.c. allowed to escape. The level of diuresis is of the same order of magnitude as that seen in clinical diabetes insipidus ; water diuresis is probably due to temporary inhibition of the secretion of

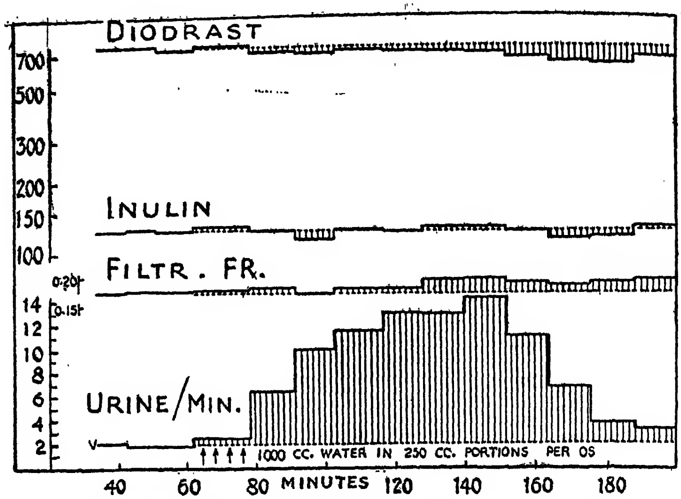


FIG. 33.—Effect of Water Drinking on Renal Dynamics. (Chassis *et al.*, *J. clin. Investig.*, 1938, *17*, 684.)

Diodrast=plasma diodrast clearance=renal plasma flow (in c.c./min.) (semi-logarithmic scale).
Inulin=inulin clearance=volume of glomerular filtrate (in c.c./min.).
Filtr. Fr.=filtration fraction=glomerular filtrate/renal plasma flow.
Urine/min. : in c.c.
At the arrows 1000 c.c. of water drunk in four 250-c.c. portions. Note that there is no significant change in renal plasma flow, volume of glomerular filtrate or filtration fraction. The urine volume increases to a peak of 14 c.c./min. from a "resting" value of 2 c.c./min.

the antidiuretic hormone of the pituitary (p. 55) reflexly from the osmoreceptors. It will be noted (Fig. 32) that the dilution of the plasma precedes the onset of diuresis. During this period it may be supposed that the circulating ADH is being destroyed and that no fresh ADH is being secreted ; as the blood ADH level falls diuresis sets in.

Ether anæsthesia inhibits water diuresis perhaps by depressing the central nervous elements involved.

Water Intoxication.—In extreme circumstances great and harmful water retention may occur. Patients with diabetes insipidus (p. 49) drink (and excrete) enormous amounts of water daily, up to 20 or even 40 L per day. If the flow of urine is restored to normal limits by injection of antidiuretic (post-pituitary) hormone *without corresponding limitation of water intake*, grave symptoms develop, *e.g.* headache, nausea, asthenia, and incoordination of movement, attributable to a great increase in the volume of all

the body fluids and a decrease of their crystalloid osmotic pressure. The condition has been reproduced experimentally. If very large amounts of water are given by stomach tube to normal dogs, retching, vomiting, and violent convulsions (resembling those of strychnine poisoning) occur at intervals, alternating with asthenia and coma. Death may take place in coma or in convulsions. If the water administration is stopped in time the animal may recover completely in a few hours. A condition of subacute poisoning can be induced in rabbits by giving daily, over a long period, an amount of water equal to the body weight. As expected (p. 128), there is œdema of the brain and a rise in the intracranial pressure. Post-mortem no structural changes are seen; the changes in the crystalloid o.p. in all the tissues plus the direct effect of the raised intracranial pressure on the brain accounts for

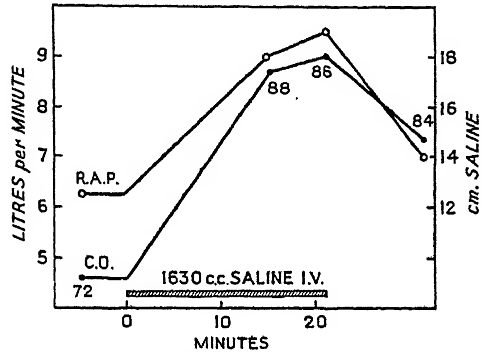


FIG. 34.—Circulatory Effects of Intravenous Injection of Isotonic Saline in Normal Man in Supine Position. (McMichael and Sharpey-Schafer, *Brit. Heart J.*, 1944, *6*, 7.)

R.A.P.: right auricular pressure in cm. saline (anterior to posterior surface of thorax), recorded by catheter in right auricle connected to manometer.
C.O.: cardiac output (in litres per minute) recorded by direct Fick method (cf. p. 279). Figures on cardiac output curve = heart rate per minute.
Inject 1630 c.c. of isotonic saline intravenously in 23 minutes. Note rise of right auricular pressure, increase in heart rate (venous reflex) and increase in cardiac output. As soon as the injection is stopped, the progressive outflow of saline from the circulation leads to a fall in right auricular pressure and in cardiac output.

most of the symptoms mentioned. These observations are important in relation to the interpretation of the symptoms of uræmia (cf. p. 77).

Effects of Injection or Ingestion of Saline Solution or Salt.—The excess sodium chloride diffuses uniformly throughout the *extracellular* fluid; *none penetrates into the intracellular fluid.* The following procedures will be considered:

(i) Rapid intravenous injection of isotonic saline.
(ii) Slow, very large, isotonic saline infusions.
(iii) Ingestion of isotonic saline.
(iv) Ingestion of excess salt (NaCl).
(v) Administration of hypotonic or hypertonic saline (pp. 126, 127).

Intravenous Injection of Isotonic Sodium Chloride Solutions.—(1) VASCULAR CHANGES.—When *very fast* rates of injection are employed the circulatory changes are very striking. On injecting 1630 c.c. in 21 minutes (Fig. 34) the right auricular pressure rises, *e.g.* from 12·5 to 19 cm. saline

owing to engorgement of the venous side of the circulation ; the heart rate is increased, *e.g.* from 72 to 88 beats per minute owing to the action of the Bainbridge reflex (p. 273) : the increased venous return produces an increased cardiac output, *e.g.* from 4·5 to 8·5 litres per minute, representing an increase in the output per beat from 60 c.c. to nearly 100 c.c. The arterial blood pressure temporarily rises. The excess fluid is first accommodated in the easily distensible veins; there is a rise of general venous pressure and capillary blood pressure corresponding to the rise of right auricular pressure. The rise of the filtering force in the capillaries drives fluid out of these vessels into the tissue spaces. At the same time, the proteins of the plasma are diluted and consequently their osmotic pressure is reduced (p. 15) ; the power of the blood to retain fluid is diminished and so exudation occurs into the tissue spaces (cf. p. 18). As the crystalloid o.p. of the extracellular fluid is

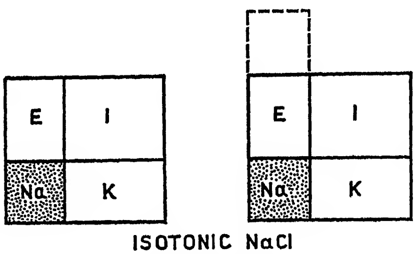


FIG. 35.—Effect of Injection or Ingestion of Isotonic Saline on Volume and Composition of Extracellular and Intracellular Fluids.

Left-hand box : *Total* areas of columns E and I represent the total volume of extracellular and intracellular fluid. Height of areas Na, K, represents the *concentration* of Na^+ and K^+.
Right-hand box : After ingestion of isotonic saline. Note the area of E is greater, *i.e.* there is an increase in the volume of extracellular fluid. There is no change in the concentration of extracellular Na^+ or of the volume or composition of the intracellular fluid.

unchanged, none of the excess fluid enters the intracellular compartment (Fig. 35).

(2) RENAL CHANGES.—There is also a greater secretion of urine ; the urine volume rises to a maximum during the second hour and then declines. In the main the diuresis is due to decreased reabsorption of water in the tubules; the mechanism is, however, obscure. As there is no decrease in the plasma crystalloid o.p. there is presumably no decrease in the secretion of pituitary antidiuretic hormone. An additional factor of somewhat greater importance than after water drinking, is the dilution of the plasma proteins which may perhaps lead to increased glomerular filtration (though this has not been convincingly demonstrated).

(3) In extreme cases some of the excess fluid may be accommodated in the *lungs*: pulmonary congestion occurs (reducing the vital capacity) and exudation into the alveoli (pulmonary œdema) may develop. Cerebrospinal fluid pressure is raised (p. 126).

(4) If isotonic saline at body temperature is injected *more gradually* (under 20 c.c. per minute) in large amounts into normal subjects there is no alteration in the arterial blood pressure or heart rate ; the excess fluid is rapidly eliminated from the plasma into the tissue spaces and is ultimately excreted in the urine, so that the plasma volume is kept about normal.

(5) The *temperature* of the fluid acts directly on the sino-auricular node (p. 236) and on the cardiac centre (p. 274); a rise of temperature quickens the heart and a fall slows it.

Very Large Slow Saline Infusions.—Fig. 36 illustrates the results of infusing 27 L of isotonic saline over 4 days (*i.e.* nearly 7 L daily). As emphasized above, as the solution is isotonic it must all be stored extracellularly until it is excreted by the kidney. The renal response though sensitive

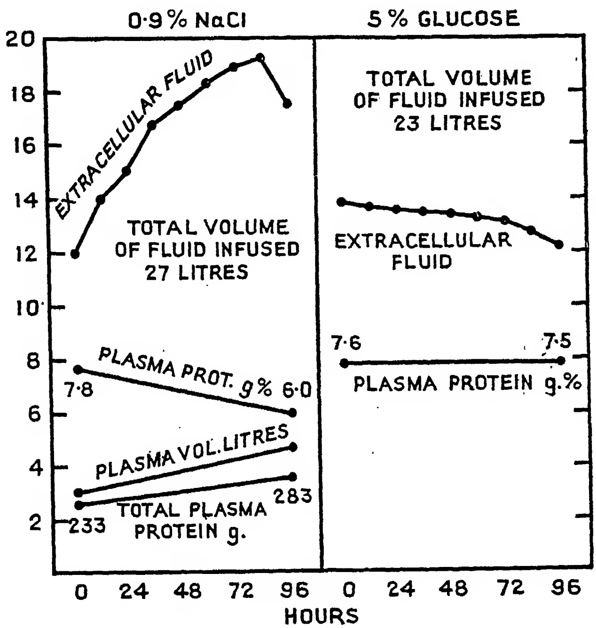


FIG. 36.—Effects of Large Slow Intravenous Infusions of Isotonic Saline and of 5% Glucose. Normal Adult. (After Gamble, *Extracellular Fluid*, Harvard University Press, 1949.)

Left-hand column: Response to isotonic saline. Note the increase in extracellular fluid volume, the increase in plasma volume and the decrease in plasma protein concentration.

Right-hand column: Response to 5% glucose. As the glucose is rapidly metabolized the effect of the infusion is equivalent to the ingestion of water. There is no increase in extracellular fluid volume or in plasma protein concentration owing to the rapid response by the kidney.

to a very small change of crystalloid o.p. (*e.g.* as after water drinking) is much more gradual and relatively inefficient when there is no change in crystalloid o.p. even when there is a large increase of total body water. Fig. 36 shows that after 84 hours there was an increase of 7 L in extracellular water, including an increase of over 1 L in the plasma volume (*i.e* of 30%); the plasma protein concentration fell (from 7·8 to 6%). The large outflow into the tissue spaces is due to the fall of plasma protein o.p. and the presumed rise of capillary blood pressure. Though the volume of urine rose to 5 L daily the kidney did not keep pace with the infusion of fluid (nor with Na^+ input). It actually responded by excreting water and extra salt in the proportions present in isotonic saline thus preserving crystalloid o.p. (there is

no change in plasma Na^+); but it obviously failed to restore total body water or plasma volume rapidly to normal.

Ingestion of Saline.—*Isotonic* saline of a composition resembling Locke's solution is fairly palatable if drunk slowly. If 1–2 L is drunk there may be *no diuresis at all* during the next 6 or 8 hours; the excess water and salt is excreted during the next few days during which period there is over-filling of the extracellular compartment; the plasma volume may be increased by *e.g.* 10% (=350 c.c.); if as much as 3 L are drunk in an hour, the maximum diuresis attained is only 300 c.c. per hour; the urinary output remains above 100 c.c. per hour for more than 24 hours. It is important to note that saline

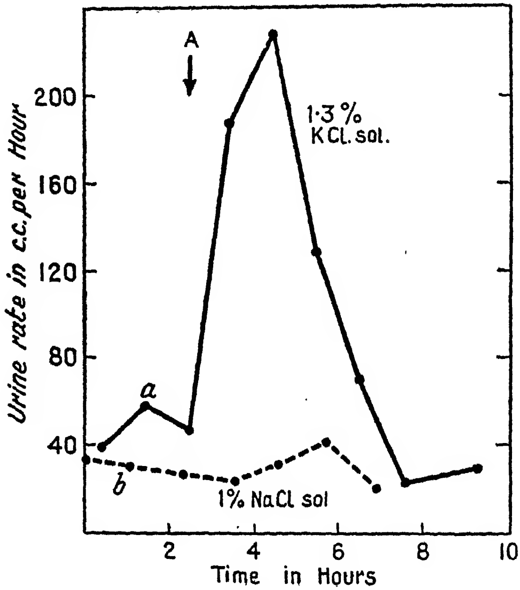


FIG. 37.—Comparison of Effects of Ingesting Isotonic (*a*) KCl Solution and (*b*) NaCl Solution on Flow of Urine. (Adolph, *Amer. J. Physiol.*, 1923, *65*.)

In (*a*) 500 c.c. of 1·3% KCl, and in (*b*) 1000 c.c. of 1% NaCl were ingested at arrow A. Note the absence of diuresis with NaCl and the marked diuresis with KCl.

is retained for many hours, when given by mouth (or per rectum), but is more rapidly eliminated if introduced directly into the circulation.

Isotonic KCl (1·3%) if given by mouth is treated quite differently from NaCl; it is rapidly excreted like urea or other waste products and gives rise to a marked diuresis (Fig. 37).

Effects of Excess Salt (NaCl).—(1) CHANGES IN BODY FLUIDS.—Suppose a strong NaCl solution is drunk or injected; the plasma crystalloid osmotic pressure rises; water is sucked from the interstitial spaces into the plasma initially increasing its volume; but at the same time the salt diffuses out into the interstitial spaces. The outflow of salt causes a secondary out-flow of water by osmotic action. Ignoring the intermediate changes just described the net result is a *uniform increase in the NaCl content of the extra-cellular fluid* without change in the relative or absolute water content of

plasma or interstitial fluid.[1] The higher osmotic pressure of extracellular fluid compared with intracellular fluid leads to a *flow of water from the cells* into the extracellular fluid. The final result is a uniform increase in crystalloid

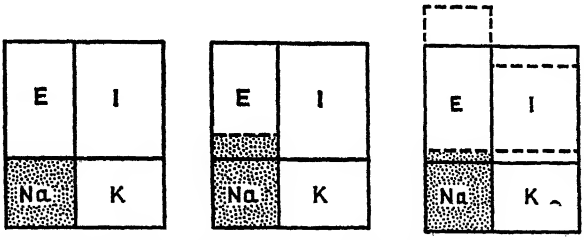


FIG. 38.—Effect on Volume and Composition of Extracellular and Intracellular Fluid of Excess NaCl.

Left-hand box: *Total* area of columns E and I represent volume of extracellular and intracellular fluid. Areas Na, K represent concentration of Na^+ and K^+ in extracellular and intracellular fluid.
Middle box: Immediate effects of NaCl excess: an increase in extracellular Na^+ concentration.
Right-hand box: Final effects of NaCl excess: an increase in extracellular fluid and a decrease in intracellular fluid volume; increase in extracellular Na^+ and increase in intracellular K^+ concentration.

concentration and osmotic pressure throughout the body fluids, but there is a *decrease in intracellular fluid* and an *increase in extracellular fluid* volume (Fig. 38).

(2) RENAL CHANGES.—In an experiment (Fig. 39) in which 28 g. of NaCl were ingested the urinary flow was increased from 30 to 120 c.c. per hour;

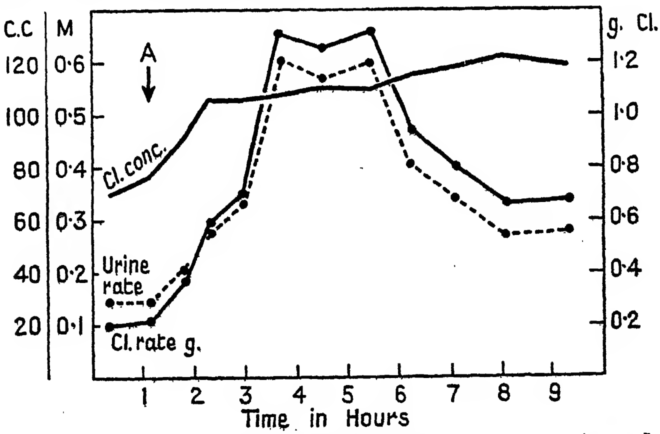


FIG. 39.—Effects on Urine of Ingesting 28 g. of NaCl. (Adolph, *Amer. J. Physiol.* 1923, *65*.)

Records from above downwards are: Cl. conc.=concentration of chloride with reference to molar concentration (M). (Molar=35·5 g. per litre); Urine rate=rate of urine formation in c.c. per hour; Cl. rate g.=rate of chloride excretion in g. per hour. At A, ingest 28 g. of NaCl.

the rate of Cl′ excretion rose from 0·2 to 1·2 g. per hour (the amount filtered per hour from the glomeruli is about 24 g.). Reabsorption of Cl′ in the tubules though somewhat less complete, is still going on on a very large (and seemingly unnecessarily large) scale. The maximal urinary NaCl concentration

[1] The small volume of water in which the salt was administered can be neglected.

occurred between the third and the twelfth hour, when it was about 0·55 molar (=3·4% NaCl, compared with the average normal of 1%). The rate of excretion of salt is thus comparatively slow.

The taking of salt induces thirst; more water is consequently drunk and is mainly retained in the body (probably owing to increased secretion of ADH (p. 54)) to dilute the retained salt and thus restore the appropriate osmotic pressure. The body weight may be temporarily increased from retention of saline. The excess fluid and salt are slowly eliminated in the course of a few *days*.

Results of Simple Sodium Chloride Deprivation.—Simple sodium chloride deficiency can be produced experimentally in various ways. The simplest is to induce severe sweating while replacing the water and NaCl lost by *water only*, excluding salt as far as possible from the diet. The changes in the body fluids are the reverse of those described for NaCl excess.

(1) CHANGES IN BODY FLUIDS.—As the NaCl content of the plasma is reduced, water initially flows out of the blood into the interstitial spaces

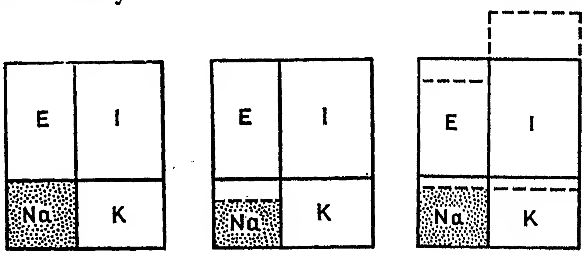


FIG. 40.—Effects of Simple NaCl Deprivation.

Left-hand box: *Total* areas of columns E and I represent volume of extracellular and intracellular fluid. Areas Na, K represent concentration of Na^+ and K^+ in extracellular and intracellular fluid.
Middle box: Immediate effects of NaCl deprivation—a decrease in extracellular Na^+ concentration.
Right-hand box: Final effects of NaCl deprivation—a decrease in extracellular and an increase in intracellular fluid volume; a decrease in extracellular Na^+ and in intracellular K^+ concentration.

(NaCl moves in the opposite direction). When equilibrium is established the volume and distribution of the extracellular fluid is unaltered but the fluid is more dilute, *i.e.* the NaCl concentration is reduced. The extracellular fluid osmotic pressure is now lower than the intracellular; fluid therefore moves from the interstitial spaces *into the cells*. The final result is a uniform decrease in crystalloid concentration and osmotic pressure throughout the body fluids, but there is an *increase in the volume of intracellular fluid* and a *decrease in the volume of extracellular fluid*; the swelling of the cells must be emphasized (Fig. 40). The decrease in extracellular water necessarily involves a decrease in plasma volume; but as this raises the plasma *protein* osmotic pressure, plasma volume is relatively better preserved than is the volume of the interstitial fluid. The concentration of Na^+ and Cl' in the plasma is decreased.

(2) RENAL CHANGES.—The kidney responds to the changes in the *electrolyte* concentration of the plasma and other body fluids by diminished excretion of Na^+ and Cl' in the urine; reabsorption of Na^+ and Cl' in the tubules becomes practically complete. A serious *disturbance of general renal function* develops which may be due to the loss of Na^+ ions.[1] Although the blood pressure is unaltered, glomerular filtration is decreased by, say, 30% and

[1] Report to Brit. Med. Res. Council, *Lancet*, 1950, ii, 509.

urea clearance is also decreased, *e.g.* to 40–80% of the control value. Considerable urea retention occurs with a resulting rise of plasma urea, *e.g.* from 30–80 mg-%; the decreased glomerular filtration may be a factor in producing this "uræmia." Water drinking does not induce the usual striking diuretic response (Fig. 41).

Na^+, Cl' *and* H^+ *ion* deprivation may be induced by continuously aspirating the gastric juice and replacing the withdrawn fluid by means of water or a glucose solution (p. 106). As under these conditions there is a greater fall of plasma Cl' than of Na^+ alkalæmia results, commonly associated with signs of renal damage, such as the appearance of albumin and casts in the urine.

Clinical Syndromes of NaCl Loss.—(1) Heat (Stoker's) Cramp.—Men working very hard in hot moist atmospheres (stokers, miners) sweat profusely; if they replace the water lost but not the NaCl, NaCl deficiency is

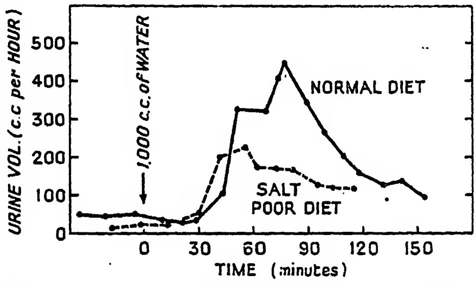


Fig. 41.—Effect of Salt-Poor Diet on Renal Response to Water Drinking. (After M'Cance and Widdowson.)

In each case 1000 c.c. of water were drunk (arrow). On a salt-poor diet the diuretic response is smaller than on a normal diet.

produced as already explained. A common symptom is widespread, intense and exceedingly painful contraction of the muscles, probably due to the harmful effect of the low NaCl content of the interstitial fluids. The cramps are prevented or markedly relieved by drinking saline (0·25%, 0·5% or even 1% solutions) or by taking salt tablets.

(2) Grave NaCl Deficiency Syndrome.—In severe cases there is lassitude, apathy, stupor, headache, weakness, as well as cramps. Later there may be nausea and vomiting (attributed to spasm of the pylorus), giddiness, fainting, pallor, and cold sweating. There is *renal failure* and nitrogenous retention in the blood. Owing to the decreased plasma volume, the blood pressure tends to fall but is kept up for a time by compensatory vasoconstriction; finally the blood pressure does fall and a "shock"-like state results. The clinical picture is thus one of general grave illness which may be variously diagnosed. A clue to the diagnosis is obtained by demonstrating the absence of Cl' in the urine (no precipitate with $AgNO_3$ solution in the presence of nitric acid), and the low plasma Cl' level. Final confirmation is obtained by the rapid improvement resulting from administration of saline solution.

Certain other points should be stressed. The normal fullness (turgor) of the skin and mucous membranes depends on the volume of *interstitial* water.

In salt deprivation the skin feels flabby and the tongue and mucous membranes are dry. *There is no sensation of thirst.*

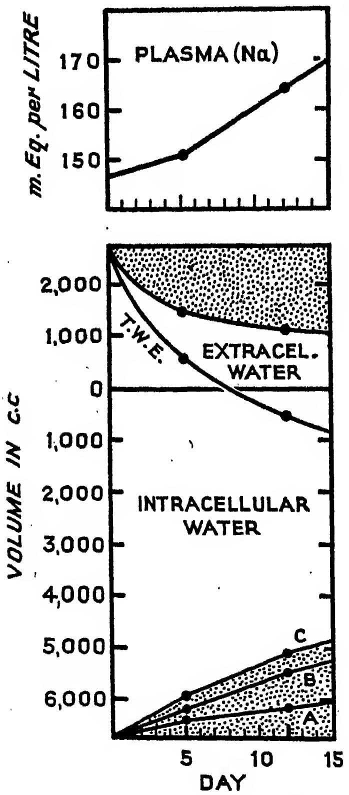


FIG. 42.—Effects on Body Fluids of Water Deprivation in Dog. (After Gamble, *Extracellular Fluid*, Harvard University Press, 1949.)

Upper record: progressive increase of plasma Na^+ concentration. A similar increase occurs throughout the extracellular fluid owing to water loss.
Lower record: Above the zero (O) line—volume of extracellular fluid; below the zero (O) line—volume of intracellular fluid.
Curve T.W.E.=total water loss=about 3500 c.c. at end of 15 days.
Upper stippled area: loss of water from extracellular compartment=1700 c.c. at end of 15 days.
Lower stippled areas A, B, C: Loss of water from intracellular compartment=about 1800 c.c. at end of 15 days.

Effect of Water Deprivation.—It will be remembered that water is constantly being lost from the body in urine, expired air, and insensible perspiration; the irreducible minimum water loss (in the absence of sweating) is about 1200 c.c. daily, representing 2·5% of the total body water. If corresponding amounts of water are not drunk (owing to water not being available, to unconsciousness, or to a patient being too weak to drink), water depletion develops leading to changes in the various compartments of body fluid which are the reverse of those described for water drinking (p. 56).

(1) CHANGES IN BODY FLUIDS.—The decrease in plasma water concentrates the plasma crystalloids; water is sucked in from the interstitial fluid and some salt diffuses out; finally the whole extracellular fluid is reduced in volume and shows a slightly increased crystalloid concentration and osmotic pressure. Water is consequently drawn out of the tissue cells. All the fluid compartments ultimately lose water and gain in crystalloid content and all end up with the same (*i.e.* slightly more than normal) crystalloid osmotic pressure. These points are well illustrated by Fig. 42, which sets out quantitative data obtained in an animal deprived of water for 15 days. Initially the water loss was borne chiefly by the *extracellular* water (almost wholly by interstitial fluid); later proportionally larger contributions were made by *intracellular* water. After 15 days of thirst the total water loss exceeded the entire initial extracellular water volume; in other words there would have been no plasma volume were it not for the contribution from the tissue cells. At 15 days, half the loss had been borne by extracellular and half by intracellular fluid. The outward flow from the tissues is chiefly due to the increased crystalloid o.p. of the extracellular fluid resulting from the dehydration; in addition there is an unexplained leak ("excretion") of K^+ out of the cells further lowering their o.p. and determining additional

water loss. The volume of urine is decreased to a minimum by increased secretion of antidiuretic hormone (p. 54) to preserve as much body water as possible; the excretion of electrolyte is raised to maximum (cf. p. 30) to minimize the very harmful increase in tonicity of the body fluids. It should be emphasized that in so-called simple water deprivation there is not only obvious water loss but also *additional electrolyte loss* for the reason just given. These facts should be borne in mind when restorative measures are planned; *chronic water loss is really water+salt loss* and must be dealt with by giving water+salt.

It is also important to remember that severe *potassium* depletion accompanies the water and salt loss found in diabetic coma and infantile diarrhœa; in these conditions one must not only treat the acidosis with alkalis and the dehydration with salines, but must make good the potassium lack by including K^+ in the transfusion fluids (cf. p. 971).

As the plasma volume falls with increasing water loss, the venous and capillary pressures fall with a resulting decrease in the filtering force which drives fluid out of the capillaries. Again, a decrease in the plasma fluid volume increases the concentration and thus the osmotic pressure of the plasma proteins. These two factors tend to keep up plasma volume at the expense of the interstitial fluid. It must be remembered that determinations of plasma volume or of hæmoglobin or of plasma protein concentration are not a reliable guide to the degree of water deprivation because the *plasma volume changes last and least* (thus maintaining the circulation); its relative constancy masks the larger water changes which are occurring extra-vascularly. When changes in plasma volume or composition are obvious, the clinical state is grave.

(2) Renal Changes in Man.—The effects of complete deprivation of fluid in man for 4 days have been studied; the diet was otherwise adequate in all constituents. The renal blood flow remained unchanged: the volume of glomerular filtrate was reduced at most by 20%. There was no decrease in plasma volume and no hæmoconcentration, presumably because the interstitial fluid reserves were called upon. The body weight decreased by about 3·5 kg. The volume of urine was reduced to 30–40 c.c. per hour; at urinary outputs of 30 c.c. per hour, or under, in man, urea, creatinine, phosphate, total nitrogenous and total non-nitrogenous solids become *maximally concentrated*. This means that larger amounts of the individual urinary constituents cannot be eliminated, should there be need to do so, unless the urinary volume is increased; conversely, any further reduction of urinary output diminishes excretion of solids and so causes their retention in the blood (cf. p. 30).

(3) Water Deprivation in Infants.—These effects of water deprivation are especially important in infants owing to the peculiarities of their renal function. *In utero*, about half the work of the kidney is done by the placenta; at birth the kidney is perhaps not fully developed functionally. The urine in infants is *hypotonic* (not hypertonic as it is normally in adults), *i.e.* an excessive volume of urine is needed (by adult standards) to eliminate a fixed amount of solid. The glomerular filtration rate is also low. Water deprivation or excess fluid loss rapidly leads to retention of urea and NaCl. In illness in children associated with sweating, diarrhœa, or vomiting, the fluid intake should be kept high and the protein intake restricted to minimize formation

of nitrogenous waste products; administration of saline should be avoided (unless there is hypochloræmia) as the kidney may fail to get rid of the extra salt.

(4) General Symptoms of Water Lack.—The patient complains of thirst and weakness; the skin is dry and inelastic; the eyes are sunken; later there may be changes in temperament; slight pyrexia may occur. The volume of urine is reduced and as the excretion of the waste products of metabolism continues, the urine is concentrated with a specific gravity up to 1040 (normal 1020). As the decreased plasma volume produces progressive circulatory failure the renal blood flow is decreased leading to renal failure, incomplete excretion of nitrogenous waste products and their consequent retention in the blood. When such symptoms are present the loss of body weight (due to water loss), is about 6% (equivalent to about 4 litres or 8% of total body water). A man who has lost 5–10 L of water is very ill; death takes place when the water loss is about 15 L.

In infants water lack is even more severely felt. The water reserves of the body are proportional to the weight and hence to the volume; the water requirements are conditioned by the metabolic rate and are hence proportional to the surface area. It follows that the smaller the individual, the greater is his surface area relative to volume, the greater is the disparity between his water requirements and his water reserves and the more quickly will water deprivation be felt.

Problem of Thirst.—The mechanism of production of thirst is not fully understood, but reference may be made to suggestive observations. Thirst, of course, is severe in conditions of water deprivation. Such thirst is not merely a mouth sensation, though the mouth is dry; thirst is said to persist after anæsthetizing the mouth but it is at once abolished by increasing the water content of the body by any route. It is common knowledge, however, that if a dry mouth is produced by arresting salivary secretion with atropine, an unpleasant sensation is produced which many would call thirst; anyway there is an urgent call for water for the mouth. In simple salt deficiency when, as explained, the *cells are swollen* (their water content is increased and the volume of the interstitial fluids is depleted), there is no thirst in man or animals. Animals given hypertonic saline suffer severe thirst and drink copiously; in this condition the *cells are shrunken* and hypertonic from loss of water, but the interstitial fluid volume is greatly increased. It would seem that the "feeling" or "affect" (as the psychologists call it) depends on the state of the body cells as a whole; how the brain is informed of this state is unknown but one may suppose that the altered state of the brain itself may contribute to the feeling of thirst. It is important, however, to recognize that thirst is *not* a guide to the *total water content* of the body though thirst is marked in simple water deprivation; it should be especially noted that the equally grave condition of salt lack gives rise to no warning sign of thirst.

Other Conditions of Altered Fluid and Ionic Balance.—The effects of fluid (+salt) loss in severe *sweating* are considered on p. 476; the effects of *vomiting* in pyloric and intestinal obstruction on p. 105; the effects of *hæmorrhage* on p. 81; of exercise on p. 53.

The rôle of the antidiuretic hormone of the pituitary in regulating crystalloid osmotic pressure is considered on p. 51.

Excretion of Urea.—The mechanism is illustrated by the following experiment. 15–30 g. of urea are ingested; rapid absorption from the intestine takes place and the blood urea rises. There is no change in renal blood flow or in glomerular filtration rate. The total urea excretion in the urine rises slowly to a peak (*e.g.* from 1·5 to 3·0 g. per hour) and then slowly declines (Fig. 43); the excess urea is gradually eliminated in the course of 8 or 10 hours. There is generally some associated *diuresis*, *e.g.* from a control level of 50 c.c. to a peak of 100 c.c. per hour; the urea concentration in the urine normally always exceeds 2% at some stage in the response and may rise to 4 or 4·5%. The mechanism of the response is straightforward: a rise of blood urea leads to the filtration of a larger *absolute* amount of urea from the glomeruli in unit time. Thus at the normal blood urea of 30 mg–% and a glomerular filtrate of 7·2 L per hour (=120 c.c. per minute), 2·2 g. of urea are filtered out per hour (depending to some extent on the volume of urine formed (p. 39)). 40–60% of the urea filtered " diffuses back " into the blood; let us assume an average " back-diffusion " of 50%; so of 2·2 g. of urea filtered, 1·1 g. is excreted in the urine. If the blood urea rises to 45 mg–% then the amount of urea filtered in the glomeruli per hour is 3·2 g. and the amount excreted (*i.e.* half) is 1·6 g. per hour. So long as the blood urea is elevated the rate of urea excretion remains raised, *i.e.* the raised blood urea " guarantees " its own excretion. Other waste products, *e.g.* SO_4'', creatinine, are excreted in a similar way.

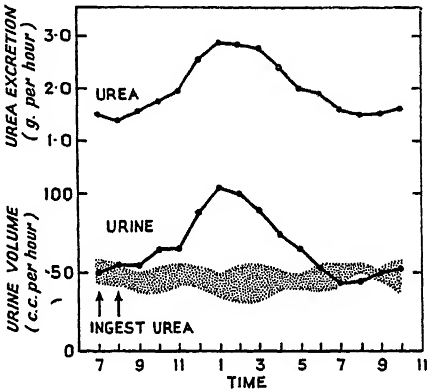


Fig. 43.—Effect of Urea Ingestion on Volume of Urine and Urea Excretion in Man. (Experiment by John Marks.)

At each arrow ingest 10 g. of urea. Note the steady progressive diuresis and gradual increase in the hourly excretion of urea.
The shaded area represents the maximum and minimum output of urine during each hour of the day in a period of 28 days on a constant water intake.

The *diuresis* produced by urea is due to the osmotic pressure exerted by the rising concentration of urea in the tubular lumen, where it interferes with water reabsorption. Urea thus retains water in the tubules for its own elimination. The maximal urea concentration in the tubules is a measure of the power of the tubules to reabsorb water against osmotic pressure resistance; it forms the basis of the *urea concentration test* which is carried out as follows:

The bladder is emptied, and 15 g. urea are given by the mouth in 100 c.c. of water. The bladder is emptied after 1, 2, and 3 hours. The first specimen may be dilute because of the diuresis induced. The second specimen should contain 2% of urea or over. If more than 120 c.c. are secreted in the second hour, the urea percentage in the third specimen is determined. Failure of the kidney to concentrate urea to 2% in any one of the three samples is evidence of renal inefficiency. *Normal kidneys* (as stated) *may concentrate urea to 4% or over.*

Because of its mechanism the diuresis induced by urea is termed *tubule*

diuresis. A similar form of diuresis takes place when *protein* is ingested or excess tissue breakdown is taking place. Larger amounts of waste products (*e.g.* urea, sulphate) are formed and their concentration in the plasma, in the glomerular filtrate, and in the tubular lumen, rises. They are excreted to a greater extent and increase the urinary volume by the same mechanisms which were considered with reference to urea ingestion. As might be expected the volume of urine varies with the *protein intake* (see Folin's Table, p. 891).

The polyuria of *diabetes mellitus* and other forms of glycosuria is likewise due to the excess sugar in the urine acting osmotically and hampering water reabsorption (p. 921).

The maximum concentration of urea in the urine is independent of that of the salts.

Rôle of Kidney in Preserving the H^+ Ion Concentration of the Blood.—This subject is discussed in detail on p. 94.

Factors Influencing Urinary Volume.—The subject has already been surveyed fully and is briefly summarized below:

(i) Water drinking (p. 57), water deprivation (p. 67) and water loss by other channels (sweating, p. 467, vomiting, p. 105).

(ii) Saline ingestion (p. 62), intravenous saline injection (p. 60), excessive salt intake (p. 63) and salt deprivation (p. 64).

(iii) Protein intake (p. 70); diuretic agents in food (*e.g.* caffeine, theophyllene and theobromine).

(iv) Muscular exercise (p. 53).

(v) *Sleep.*—During sleep the urinary output is decreased; this is attributed to a depressed state of the circulation, to lowered metabolism and to lack of water intake.

(vi) Emotion may inhibit (p. 53) or increase the flow of urine; it may also increase the frequency of micturition (p. 771).

(vii) *Rôle of posterior pituitary.*—By means of its antidiuretic hormone the neural division of the pituitary controls the daily volume of urine; it acts directly on the tubular epithelium increasing the reabsorption of water (p. 51).

(viii) *Other Endocrines.*—Injection of adrenaline (p. 728); adrenal cortex (p. 955); thyroid (p. 51); anterior pituitary (p. 51); parathormone (p. 1007); diabetes mellitus (p. 921).

(ix) *Other Factors.*—Hæmorrhage (p. 83); traumatic shock (pp. 340, 342); action of hypertensin (p. 351); essential hypertension (p. 354); acidæmia (Figs. 54, A, 55); alkalæmia (Figs. 54, B, 56); renal disease (pp. 71 *et seq.*).

KIDNEY FUNCTION IN DISEASE[1]

Tests of Renal Function.—1. Examination of Blood and Urine.—The *urine* should be examined with respect to volume, specific gravity, reaction, concentration, absolute amounts of the various normal urinary constituents, and presence of abnormal constituents like albumin, red cells, or casts, and the *blood* with respect to the concentration of the constituents excreted and thus regulated by the kidney including H^+ ion concentration.

[1] Peters and Van Slyke, *Quantitative Clinical Chemistry, Interpretations*, 1931. Fishberg *Hypertension and Nephritis*, 1939.

(1) Specific Gravity and Urea Concentration Test.—The power of the kidney to absorb water and do osmotic work is gauged by the *specific gravity* of the urine (p. 29) and more precisely by the *urea concentration test* (p. 69) which also indicates the ability of the tubular epithelium to resist " back-diffusion."

(2) Significance of Blood Values.—In considering the effect of renal failure on the level of the various constituents of the plasma the following facts must be borne in mind.

(i) *Urea Excretion and Blood Urea Level.*—As explained on p. 69, urea is eliminated from the body exclusively by glomerular filtration; the amount in g. of urea filtered into Bowman's capsule=glomerular filtrate (in L) × plasma urea concentration (in g. per L). Of the amount so filtered about half returns to the blood by " back-diffusion " and the other half is excreted in the urine. The table below shows the minimal blood urea level necessary to eliminate 9, 18, and 36 g. of urea daily (resulting from low, modest, and high protein intake) at normal and subnormal levels of glomerular filtration.

Glomerular Filtrate		Blood Urea Leve (in mg./100 c.c.) necessary to excrete		
Volume 24 hours.	Fraction of Normal. 1·0=Normal.	9 g. of Urea.	18 g. of Urea.	36 g. of Urea.
180 L	1·0	10 mg–%	20 mg–%	40 mg–%
90 ,,	0·5	20 ,,	40 ,,	80 ,,
45 ,,	0·25	40 ,,	80 ,,	160 ,,
22·5 ,,	0·125	80 ,,	160 ,,	320 ,,

At a normal level of glomerular filtration (180 L), 36 g. of urea can be excreted at a blood urea concentration of 40 mg–%. Suppose the glomerular filtrate falls to half-volume (90 L); the amount of urea excreted is similarly halved; urea is thus " retained " in the blood and its concentration there rises. When the blood urea has risen to double its original level (80 mg–%) the kidneys again excrete 36 g. (at the reduced glomerular filtration rate) and equilibrium is re-established. On the other hand if only 18 g. of urea have to be excreted, when the glomerular filtrate falls to half, a blood urea level of 40 mg–% is adequate to effect full excretion. If the urea to be excreted is only 9 g., a blood urea level of 40 mg–% is adequate even when the glomerular filtrate is reduced to *one quarter* (45 L). A reduction of glomerular filtration (with constant urea production) leads to a rise of urea concentration in the blood (and throughout the body fluids), *i.e.* there is urea retention. But this rise of blood urea is eminently " useful " because it automatically leads to a larger urea excretion; at an appropriately raised level of blood urea a balance between excretion and production can always be struck. As far as is known a rise of blood urea *per se* is harmless (cf. p. 74) and as urea is freely diffusible it does not disturb the balance between the various compartments of body fluid.

A raised blood urea level must be considered in the light of these facts. In renal failure the raised blood urea can always be lowered to some extent by decreasing the protein intake, the lower blood urea level then being adequate to eliminate the smaller amounts of urea which are formed.

(ii) The Van Slyke urea clearance test described on p. 39 and summarized in the expression $[C_s = \frac{U_m}{P} \times 100]$ gives a numerical value to the relationship between the amount of urea excreted (U_m) and the plasma urea level (P) necessary to excrete it. As P is the denominator, any increase in plasma urea level needed to excrete a given amount of urea lowers the clearance value; the fall of clearance value is an index of renal insufficiency.

(iii) The *normal blood urea* value is generally 25–40 mg. per 100 c.c.; lower values are, however, not uncommon and are the rule in pregnancy. In health the blood urea will obviously fluctuate with the protein intake. A rise of blood urea is a sensitive index of renal failure if the protein intake is high; on a low protein intake considerable renal failure may be present without the blood urea rising above the upper limit of the normal.

(iv) *Blood Urea in Non-Renal Disease.*—A raised blood urea (in relation to the protein intake) probably represents decreased glomerular filtration, but the latter need not necessarily be due to renal disease. Any condition decreasing glomerular blood flow or glomerular pressure decreases glomerular filtration. The blood urea may thus be raised in hæmorrhage (p. 84), shock (p. 343), conditions of water deprivation (p. 67), salt deprivation (p. 65), or combined water and salt loss (*e.g.* intestinal obstruction, p. 108), over-treatment with alkali (used "therapeutically"), in adrenal insufficiency (p. 955), in parathyroid hypercalcæmia (p. 1007), circulatory failure, or as a result of reflex constriction of the renal vessels (p. 27).

(v) *Blood Non-Protein Nitrogen (N.P.N.).*—The term non-protein nitrogen (N.P.N.) refers to the weight of nitrogen in the non-protein nitrogenous constituents of the plasma (*i.e.* such substances as urea, mixed amino-acids, uric acid, creatinine); it is expressed in mg. per 100 c.c. and is normally about 20–40 mg. [Note that blood urea is expressed as the *total* weight of urea; if the blood urea is 40 mg., then the blood *urea nitrogen*[1] is $\frac{28}{60} \times 40 =$ 18 mg.; approximately half the blood N.P.N. is thus due to urea N.]. The *plasma* N.P.N. is 18–30 mg. per 100 c.c. The determination of the blood urea and the N.P.N. gives an indirect indication of the concentration of the nitrogenous constituents other than urea. The blood urea and N.P.N. usually rise and fall together and have the same significance.

The principal non-protein nitrogenous constituents of the plasma other than urea are *creatinine* and *uric acid.*

Blood Creatinine.—The upper normal limit of blood creatinine is 2 mg. per 100 c.c. It is thought likely that to excrete the creatinine normally made in the body approximately the *full* normal glomerular filtration volume is necessary. As creatinine is mainly formed in the tissues at a fairly steady rate its blood level is normally very constant. Any decrease in glomerular filtration would then be immediately reflected in a rise of blood creatinine.

[1] Urea N/Total Urea $= 2N/CO(NH_2)_2 = 28/60$.

Values of 3–5 mg–% indicate considerable glomerular failure, while values over 5 mg–% are of grave omen.

Blood Uric Acid.—The normal range is 0·7–3·7 mg. per 100 c.c. (average 2 mg.). As the urinary uric acid is partly derived from purines in the food (p. 897) the blood level will be influenced by the diet, especially in renal failure. It is claimed (but not explained) that uric acid is the first nitrogenous constituent to be retained in nephritis; in uræmia values as high as 27 mg–% have been recorded. Blood uric acid is, however, also raised in *gout* (p. 898) and in conditions associated with excessive destruction of the nuclei of white blood corpuscles (leukæmia, pneumonia).

(vi) *Plasma Na^+ Cl', K^+, and Phosphate.*—With these *threshold* substances the problem of their excretion by the kidney is quite different from that of urea and the other no-threshold substances (see Table).

	Na^+	Cl'	K^+	Phosphate
Plasma concentration (mg./100 c.c.) . .	330 (mg.)	365	17	3
Total filtered in 170 L of glomerular filtrate in g. (24 hours) .	560 (g.)	620	20	5·1
Amount normally excreted in urine in g. (24 hours) . .	5 (g.)	9	2·2	1·2

In the case of Na^+ and Cl' about 99% of the amount filtered is reabsorbed, 1–2% being excreted. Reabsorption is adjusted to maintain the appropriate level in the plasma, and is diminished whenever Na^+ or Cl' tend to accumulate. Examination of the data in the Table shows that the volume of glomerular filtrate could be reduced to *one-fiftieth* of normal while permitting full excretion of these substances. A rise of Na^+ or Cl' level in the blood can *never* be the result of *reduced* glomerular filtration but only of virtually *complete cessation* of such filtration (or of *pathologically increased reabsorption*). The K^+ intake could be excreted at *one-fifth* the normal filtration rate and HPO_4'' at *one-quarter*. Retention of phosphate and perhaps of K^+ (rise in their blood level) does occur in renal failure. [The effect of parathormone on phosphate excretion is considered on p. 1003; for effect of adrenal cortex on Na^+ and K^+ excretion see p. 945.]

2. Detailed Study of Renal Dynamics.—In more refined analyses of renal function the following data should be obtained: (*a*) *renal plasma flow* and *total blood flow* (diodrast clearance value (p. 36)); (*b*) *glomerular filtration volume* (inulin clearance value (p. 35)); (*c*) *filtration fraction* (p. 27); *tubular excretory maximum* (T_m), *i.e.* maximal power of the renal tubules to excrete specific substance, *e.g.* diodrast (p. 40).

Extirpation of the Kidneys. — Latent (Asthenic) Uræmia.[1]— Removal of *one* kidney results only in a transient increase in the non-protein nitrogenous compounds (N.P.N.) of the blood. If three-quarters of the renal tissue is removed, recovery may take place with little change in the blood urea and N.P.N. if the protein intake is modest (cf. p. 71), but on high protein diets the blood nitrogen rises.

If *both* kidneys are removed, or if there is obstruction of both ureters or

[1] Harrison and Mason, *Medicine*, 1937, *16*, 1.

of the renal tubules, a clinical syndrome develops to which the term *latent* (*asthenic*) *uræmia* is applied (cf. p. 183). There is a great increase in the non-protein nitrogenous constituents (N.P.N) of the blood. The rise in blood *urea* may sometimes be to as much as 900 mg. per 100 c.c. *Ammonia* does not appear in the blood (it is *normally absent* from blood) as it is made by the kidneys. The patient looks ill and cachectic, anæmia develops and body temperature becomes subnormal. There is a yellowish-grey discoloration of the face and hands, and eruptions (vesicular, petechial, or hæmorrhagic) appear on the skin. The mental state is one of apathy, listlessness, and ready fatiguability; the patient moves and speaks slowly, dozes frequently, but answers intelligently when roused. The tendon reflexes are initially exaggerated and later depressed; muscular twitching is common but *convulsions are rare.* There is a good deal of gastro-intestinal derangement, consisting of vomiting (which may be copious even when the changes in the blood are still slight), colitis, and diarrhœa. Respiration becomes deep and sighing. There is *little if any rise of blood pressure,* even terminally. Deep coma sets in a few hours before death, which occurs after 5–7 days.

MECHANISM.—There is uncertainty as to the mode of causation of the symptoms of latent uræmia.

(i) There is surprisingly little evidence that the accumulation of the end products of protein metabolism is harmful. It is true that if *extremely large* amounts of *urea* are given by mouth to normal people they prove toxic; but an important factor may then be the increased diuresis and resulting dehydration. Nitrogenous retention *without water loss* can be produced by implanting the ureters into the small intestine; the nitrogenous bodies are reabsorbed into the blood. When their concentration rises markedly, no obvious symptoms are set up. Again, there is no close relationship in cases of latent uræmia between the blood urea level and the degree of intoxication.

(ii) Urea (in uræmia) is present in high concentrations in the *intestinal mucosa*; it is suggested that it may decompose locally to form ammonia which is highly irritating and may be responsible for the gastro-intestinal disturbance. A *urease* (capable of splitting urea and releasing NH_3) has been demonstrated in normal gastric mucosa.

(iii) Large doses of creatinine are not toxic.

(iv) There is little change in the blood reaction.

(v) The serum Cl' and Na^+ may rise because of retention if the secretion of urine ceases, or fall through loss in the vomit or fæces. Disturbances in the electrolyte balance may produce serious symptoms (cf. p. 65).

(vi) Serum K^+ and phosphate increase as a result of "retention." The rise in phosphate may secondarily lower serum Ca^{++} level. Changes in K^+ and Ca^{++} balance are known to affect the excitability of nerve fibres, nerve centres, and skeletal muscle (pp. 520, 523, 1004).

(vii) Finally some unknown organic metabolites may be a factor in causing the coma.

Nephritis.—The classification employed here is that of Ellis, Evans, and Wilson [1] which is based on a follow-up of about 600 cases over a period of 20 years; 200 of these cases were examined histologically after death. They divide nephritis into two main types, *i.e.* Type 1 and Type 2.

[1] Ellis, *Lancet*, 1942, i, 34, 72. Hadfield and Garrod, *Recent Advances in Pathology*, 5th Edn., London, 1947.

(1) **Type 1 Nephritis** (*diffuse nephritis; glomerulonephritis; inflammatory Bright's disease*).—This condition affects mainly children and young adults. A history of acute infection, usually of the upper respiratory tract is generally obtained. The onset is abrupt and accompanied by malaise, vomiting, and headache; gross hæmaturia is usually found at this stage. Œdema is present at the onset affecting the face and ankles; it is not progressive and is of short duration. Hypertension of varying degree is an early sign in most cases. Over 80% of the cases make a complete recovery; the rest either die in the acute stage or pass into a state of chronic nephritis.

In fatal cases the *glomeruli* throughout the kidneys are affected by a specific inflammatory process. Initially there is increased glomerular permeability as shown by the passage of blood or albumin in the urine; histologically the capsular space is filled with exuded leucocytes. The capillary endothelium of the glomeruli and the epithelium of Bowman's capsule proliferate and obstruct the blood flow through the glomeruli. This inflammatory renal ischæmia (like experimental renal ischæmia (p. 345)) may be supposed to cause hypertension by release of excess renin and consequent production of hypertensin (p. 349). In cases that do not resolve quickly the hypertension in its turn damages the blood vessels, with a special predilection for the afferent glomerular arterioles. The ischæmia of the kidney is aggravated, more renin and hypertensin are formed, and a *hypertensive vicious circle* is set up which (as explained on p. 357) leads to progressive destruction of the kidneys and death from uræmia, heart failure, or vascular accidents.

The renal hæmodynamics have not been adequately studied in acute cases of Type 1 nephritis. It can be assumed from the pathological changes that both the renal blood flow and the glomerular filtrate volume are reduced, leading to failure of renal excretory activity. The volume of urine is reduced, *e.g.* to 100 c.c. or less; the urine contains blood, albumin which has leaked out of the plasma as it passes through the glomeruli, and casts from the injured tubules. Water, electrolytes, and waste products are retained; the plasma volume is increased; the blood urea and non-protein nitrogen rise; the plasma protein concentration falls. It has been observed that in cases in which the blood pressure is normal or only slightly raised the volume of urine is very small or anuria develops; the moderate hypertension commonly present thus has some compensatory value in driving blood through the affected glomeruli. If the hypertension is severe (over 170 mm. Hg) cerebral changes occur described as *hypertensive encephalopathy* which are probably the result of abnormalities in the blood supply to the brain; the symptoms are severe headache, vomiting, excitement, convulsions, blindness, and coma.

(2) **Type 2 Nephritis** (*chronic parenchymatous nephritis; nephrosis*).—The age distribution is wider than in Type 1, most of the cases occurring in adults. The onset is insidious without any history of an acute infection. Probably the first indication of the disease is *albuminuria*; there is no hæmaturia at the onset. Patients usually first seek advice because of the development of *œdema*, which is invariably present and is progressive, persistent, and often very severe. About 95% of the patients die of the disease.

In the œdematous phase the distinctive *glomerular* lesion is a hyaline thickening of the basement membrane of the wall of the capillary loops;

this anatomical change is associated with greatly increased capillary permeability to protein which leads to the loss of albumin in the urine. In this phase the renal *tubules* are characteristically affected: the lining cells contain large amounts of anisotropic lipide material. The œdematous phase may persist for years and the patients are very liable to die of secondary infection. The glomerular change is relentlessly progressive and sooner or later the glomerular capillaries become obliterated and hyalinized. The result is renal ischæmia with resulting hypertension and renal failure.

The glomerular change in Type 2 nephritis (*i.e.* increased permeability with adequate blood flow) accounts for the albuminuria which in its turn is indirectly the main cause of the œdema (p. 112). The functional results of the early tubular changes are unknown.

Other Causes of Renal Injury.—As explained on p. 356 the commonest cause of renal damage is essential hypertension. The kidney is also involved in a variety of miscellaneous disorders like pyelonephritis or polycystic kidney.

The end result in all forms of renal damage, if the patient survives long enough, is progressive obliteration of the glomeruli, atrophy of the corresponding tubules; renal ischæmia; renal failure; hypertension and its usual sequelæ; this condition may be called *chronic nephritis.*

Chronic Nephritis—In this phase of the disease patients generally show varying degrees of reduction in renal blood flow and in the volume of the glomerular filtrate, depending on the severity of the case. Nitrogenous retention (raised blood urea and N.P.N.) consequently develops. Damage to the renal tubules impairs the reabsorption of water; there is polyuria and the urine has a low specific gravity. Failure to excrete acid radicals (phosphate and SO_4'') may lead to acidæmia; this condition is aggravated by the diminished power of the kidney to make NH_3 to neutralize the excess acid radicals and by the decreased reabsorption of $BHCO_3$. Failure to reabsorb the normal amounts of Na^+ and Cl' may cause some salt deprivation. The failure of the NH_3 mechanism involves *drainage of fixed base from the body.* As base is being lost there is no rationale for the conventional low salt diet in this type of nephritis. An *ample salt intake* should be provided to preserve the electrolyte pattern of the plasma. The plasma changes are thus complex and variable. Examination of renal function by modern methods in chronic nephritis indicates that patients show varying degrees of reduction of renal blood flow and of the volume of glomerular filtrate, depending on the severity of the case.

Sthenic Uræmia.—During the acute phase of Type 1 nephritis or in the terminal phases of chronic nephritis or essential hypertension a syndrome may develop known as *sthenic uræmia*, characterized by marked nitrogenous retention, dyspnœa, marked hypertension, often epileptiform convulsions, and gastro-intestinal disturbances. These findings are due partly to failure of renal excretory activity and partly to circulatory disorders.

(1) Blood Changes.—These are due to renal failure, and the changes in fluid and electrolyte balance resulting from diarrhœa, vomiting, and the forms of treatment employed.

(i) There is an increase of the non-protein nitrogen (N.P.N.) from the normal level of 25–40 mg–% to 50, 100, or even 300 mg–%; the blood urea may be 200 mg–%, the uric acid 5–27 mg–%, and the creatinine 2–40 mg–%. Exceptionally the level of blood nitrogen may be little altered.

(ii) An increase in serum phosphate and a (consequent) fall in the serum calcium (to 6 mg-%) have been noted in some cases; this may be the cause of the twitching that occurs, *i.e.* it is analogous to that seen in *tetany* (p. 1007).

(iii) In some cases of uræmia there is a considerable "metabolic" acidæmia owing to the renal disturbances described on p. 76; there is consequently great hyperpnœa, and a resulting compensatory fall in the alveolar and arterial CO_2 tension (p. 456). The acidæmia may give rise to convulsions and coma, and further damage the renal tissue. If alkali is given in sufficient doses to restore the normal blood reaction, the convulsions may cease, consciousness is restored, and there is improved renal function, as shown by a diminished amount of albumin and casts in the urine. The improvement is of necessity only temporary.

(2) Respiratory Changes.—As a result of the poisoned and depressed condition of the respiratory centre, Cheyne-Stokes breathing (p. 464) and paroxysmal dyspnœa (p. 463), especially at night, may be present.

(3) Gastro-Intestinal Changes.—Vomiting and diarrhœa are common. Vomiting may in part be of cerebral origin. Both disturbances may be due to irritation of the mucous membrane of the bowel from the excretion into it of some hypothetical poisonous body or the formation of ammonia from retained urea (cf. p. 74). Ulceration of the bowel (colon, lower ileum) is often present, and may be associated with bloody stools.

(4) Hypertension and Convulsions.—The hypertension in sthenic uræmia is not due to the retention of substances normally excreted in the urine as it does not occur after double nephrectomy or blocking of the ureters (p. 74). The hypertension is due to neurogenic or humoral factors, *e.g.* increased sympathetic activity or pressor substances released by the kidneys or other organs.

Convulsions like those seen in sthenic uræmia may occur without signs of renal failure in essential hypertension; the syndrome is known as *hypertensive encephalopathy*. The convulsions are attributed to cerebral anæmia resulting from spasm of the cerebral arterioles or obstruction of these vessels by small thrombi.

Another factor may play a part in the renal cases, namely *cerebral œdema* and *raised intracranial pressure*. Attention has been drawn (p. 59) to the resemblance between the symptoms of water intoxication and those of sthenic uræmia. Lumbar puncture in uræmia frequently reveals the cerebrospinal fluid to be under increased pressure; after removal of the excess fluid temporary improvement may occur. One can readily understand how cerebral œdema may cause headache, vomiting, loss of sight, convulsive movements, and coma (p. 125). It may also be supposed that hypertension may sometimes develop secondarily to the increased intracranial pressure; it would then represent an attempt to maintain an adequate cerebral circulation.

Complete renal failure causes death in about 10 days. In some cases of acute failure there is reason to believe that if the patient could be kept alive sufficiently long the kidney might recover its functions; for this reason attempts have been made to devise a temporary artificial substitute for the kidney. It is difficult to assess the value of these somewhat heroic methods as spontaneous recovery sometimes occurs after a number of days without such treatment. The two principal methods used are (i) the "artificial kidney"; (ii) peritoneal lavage.

Artificial Kidney.—The apparatus consists essentially of 30–45 metres of cellophane tubing spirally wound round a cylinder which is rotated in a tank of warm "rinsing" fluid consisting of a modified Ringer's solution. The patient is heparinized and his blood flows from a cannula inserted into the radial artery through the cellophane tube whence it is pumped back into a vein. In this "kidney" exchange of diffusible crystalloid substances takes place between the plasma and the surrounding rinsing fluid. As the latter is free from nitrogenous waste products, these diffuse out of the plasma

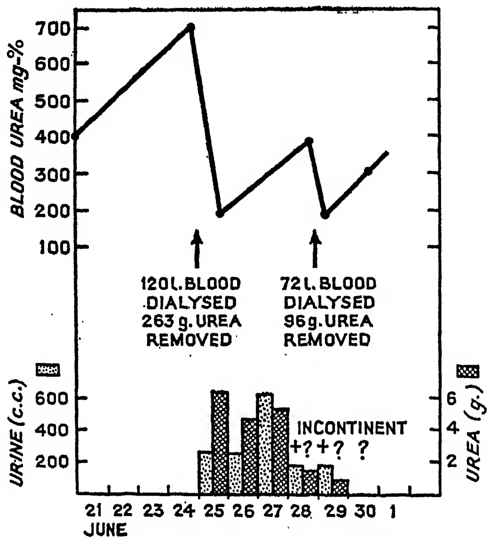


FIG. 44.—Results of Treatment with Artificial Kidney.

Upper record: Blood urea in mg. per 100 c.c.
Lower records: Volume of urine in c.c. and amount of urea in grammes excreted each day.
The records are from a man aged 38 who had shown signs of glomerulonephritis for 3 weeks. The blood urea on admission was 700 mg-%; blood pressure: systolic 220, diastolic 120 mm. Hg.
In the first dialysis 120 litres of blood were circulated through the artificial kidney; 263 g. of urea were thus removed and the blood urea fell to 190 mg-%.
During this period the *daily* excretion of urea in the urine was only about 6 g.
The blood urea rose again to 385 mg-%. A second dialysis of 72 litres of blood led to the removal of 96 g. of urea. The patient died of heart failure. (W. J. Kolff, *New Ways of Treating Uræmia*, Churchill, London, 1947.)

and are thus "excreted"; by suitably modifying the composition of the rinsing fluid the plasma can be made to take up or lose such constituents as water, Na^+, K^+, Cl', or H^+ ions. Though the blood flow through the machine is 100–200 c.c. per minute compared with a normal renal blood flow of 1200 c.c., it can "clear" 80–150 c.c. of plasma of its urea content, compared with a normal renal urea clearance of 75 c.c. Figs. 44 and 45 show that very large amounts of urea can be eliminated by the artificial kidney with a corresponding lowering of the initially high blood urea level; other waste products are simultaneously got rid of, *e.g.* the blood creatinine may fall from 5·4 to 2·3 mg–% and the blood uric acid from 9·7 to 2·3 mg–%. Kolff[1] recommends that the use of the apparatus be considered when the blood urea exceeds

[1] Kolff, *New Ways of Treating Uræmia*, London, 1947; *Cleveland Clin. Quart.*, 1950, *17*, 216.

350 mg–%. The great limitation of the machine is that it is inanimate and lacks the power of the normal kidney of sensitively responding to changes in blood composition in such a manner as to preserve the constancy of the *milieu intérieur*. With the machine all kinds of derangements of fluid and electrolyte balance may occur with all their undesirable consequences. But some lives have been saved by its use.

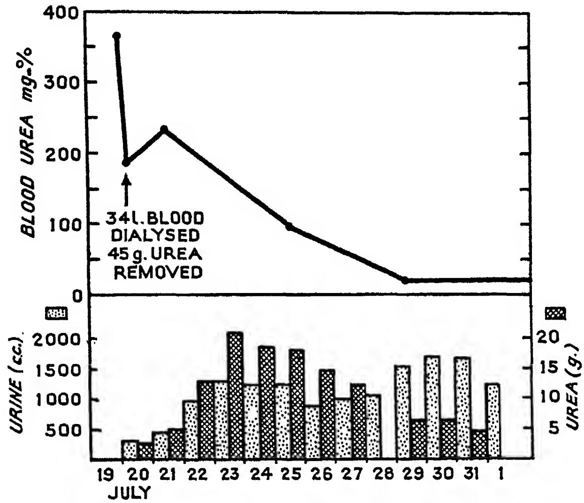


FIG. 45.—Results of Treatment with Artificial Kidney.

Upper record: Blood urea in mg. per 100 c.c.
Lower records: Volume of urine in c.c. and amount of urea in grammes, excreted each day.
The records are from a girl aged 13 who developed acute glomerulonephritis. On admission she had œdema of the lungs, bronchopneumonia, cyanosis, cardiac enlargement, and an enlarged tender liver. She was very ill, restless and disorientated. Blood pressure: systolic 120, diastolic 60 mm. Hg. There had been anuria for a week before admission.
Blood urea on admission, 365 mg–%. 34 L of the patient's blood were dialysed through the artificial kidney in 4½ hours; during this period 45 g. of urea were removed. The blood urea fell from 265 to 140 mg–%. The flow of urine restarted during the dialysis and progressively increased. The patient recovered. (W. J. Kolff, *New Ways of Treating Uræmia*, Churchill, London, 1947.)

PERITONEAL LAVAGE.—The rinsing fluid is perfused through the peritoneal cavity at the rate of 1 litre per minute, and diffusion takes place between it and the intestinal and peritoneal capillaries. This method is much less efficient than the artificial kidney, the urea clearance for example being only 15 c.c. per minute; but it has the merit of relative simplicity. Even with this procedure control of the electrolyte balance by means of frequent blood determinations is essential.

PLASMA VOLUME. CELL VOLUME. TOTAL BLOOD VOLUME. REGULATION OF BLOOD VOLUME[1]

The *measurement* of plasma volume, cell volume, and total blood volume is described on pp. 9 *et seq.*

Plasma volume regulation is part of the general problem of body water regulation. *Cell* volume depends on the number and size of the cells; as

[1] Rowntree and Brown, *Volume of Blood and Plasma in Health and Disease*, Phila., 1929.

the vast majority are erythrocytes the cell volume depends on the activity of the blood-forming system (red marrow) and the blood-destroying system (macrophages). Plasma volume and cell volume may vary independently of one another and are influenced by quite distinct sets of factors. In abnormal states large variations in corpuscular volume occur (due to anæmia or polycythæmia); these are more compatible with life than changes of similar magnitude in the plasma volume. The state of the *total* blood volume depends on the accuracy or otherwise with which its constituent parts—plasma volume and cell volume—are regulated.

Normal Blood Volume.—The volume of the blood in normal subjects (male and female) is 70–100 c.c. per kg. body weight; that of the plasma is 40–60 c.c. per kg. (average 50 c.c. in males, 48 c.c. in females). The volumes are more closely related to surface area; per square metre of body surface (cf. p. 378) the volume of the blood is 2500–4000 c.c. and of the plasma 1400–2500 c.c. (average 1·8 litres in males and 1·6 litres in females). The total blood volume is about one-eleventh of the body weight, or 6 litres in an average adult.

The blood and plasma volumes vary with age; before puberty, the volumes in relation to body weight are below adult standards. The blood volume increases during pregnancy and falls again after delivery; this fall is only partly accounted for by the hæmorrhage associated with parturition.

Distribution of Blood in Different Organs.—Few data on this subject are available for man. The *two legs and arms* together contain at rest about 20% of the blood volume. If occluding cuffs are applied to three of the limbs and distended at diastolic blood pressure, the arterial inflow continues but the venous return is stopped; maximal venous engorgement of the limbs takes place. The *additional* volume of blood in the limbs may be 900 c.c. This procedure depletes the general circulation of blood and produces the effects of a venesection of equal size; it may be employed as an emergency measure in cases of congestive heart failure to decrease the central venous pressure and the load on the heart (Fig. 159, p. 275).

The *lung* vessels in man contain about one litre of blood. This blood may serve as a reservoir and be discharged in part into the systemic circulation during emergencies; thus after venesection to the extent of 380 c.c., the vital capacity may increase by 150 c.c. and the total lung volume by 180 c.c. owing to a decrease in the capacity of the pulmonary vascular bed. Conversely, in circulatory plethora excess fluid is stored in congested pulmonary vessels.

Regulation of Plasma Volume.—This is effected in two main ways:

(i) By appropriate adjustments between the plasma and the rest of the body fluid;

(ii) by adjustments in the volume of urine secreted.

Some relevant general observations on the regulation of the volume of fluid in the different compartments, fully discussed elsewhere (pp. 55–67), may usefully be summarized here.

(1) Simple *excess water* is stored uniformly throughout all the fluid compartments, with an ultimate uniform fall in the crystalloid osmotic pressure of all the body fluids.

(2) Simple *water loss* leads to a uniform decrease of water content and an increase in crystalloid osmotic pressure in all the fluid compartments.

It should be noted that in simple water excess or loss, the plasma volume

is "buffered" (protected) by all the rest of the body fluid. Again, as the secretion of the antidiuretic hormone is sensitively regulated by the crystalloid osmotic pressure, the response of the kidney to simple water excess or loss is prompt and effective, so not only is the volume of the plasma preserved but also that of the other body fluids (p. 53).

(3) Simple *salt* (*NaCl*) *excess* raises the crystalloid osmotic pressure of the plasma and the interstitial fluid; consequently fluid is withdrawn from the intracellular into the extracellular compartment. Extracellular fluid including plasma volume increases, while intracellular fluid volume decreases. In other words, the normal distribution of fluid in the three compartments is "sacrificed" in the interest of minimizing the rise of crystalloid concentration in the extracellular fluid (p. 62).

(4) In simple *salt* (*NaCl*) *deficiency* the reverse changes occur (p. 64).

The renal responses in (3) and (4) are considered on pp. 63, 64.

(5) Administration of an *isosmotic* (*isotonic*) *NaCl solution* leads to an increase in the extracellular fluid volume including that of the plasma. As there is no change in the crystalloid osmotic pressure of the plasma there is no movement of fluid into the intracellular compartment which in this condition does *not* help to protect the extracellular fluid volume (p. 60). The excess plasma volume is transferred to the interstitial compartment, not by crystalloid osmotic pressure changes (for none occur) but owing to an *increase in capillary blood pressure* and a decrease in *plasma protein osmotic pressure* (p. 15). The renal changes are considered on p. 60.

The preservation of the relative constancy of the plasma volume in a variety of conditions is discussed fully elsewhere as follows:

(i) Effects of water drinking, p. 55.
(ii) Effects of water deprivation, p. 66.
(iii) Effects of salt excess, p. 62.
(iv) Effects of salt deprivation, p. 64.
(v) Effects of intravenous injection of isotonic saline at different rates, pp. 59, 61.
(vi) Effects of drinking isotonic saline, p. 62.
(vii) Effects of ingesting or injecting hypotonic or hypertonic saline, pp. 126, 127.

The interchange of fluid between the plasma and the tissues at rest and during activity is considered in detail on pp. 18–21.

Hæmorrhage.—The effects of hæmorrhage depend not only on the *volume* of blood which is lost, but also on the *rate* at which this loss occurs. The compensatory powers of the body are very great, but they take time to develop their maximum effect, so that a sudden large hæmorrhage is more serious than a greater loss spread out over a more prolonged period.

Immediate Effects of Severe Hæmorrhage.—The effects of a large acute hæmorrhage in man have been studied by withdrawing 15–20% of the blood volume (750–1200 c.c.) in 6–13 minutes under carefully controlled conditions in professional donors who had been repeatedly bled and were thoroughly familiar with the effects of venesection.[1] As the bleeding proceeded the face turned pale, and the hands became cool and sweaty. The systolic blood pressure fell by about 10 mm. Hg, the diastolic pressure was

[1] Ebert, Stead and Gibson, *Arch. int. Med.*, 1941, 68, 583.

unchanged, and the heart rate rose by 15–30 beats per minute; the venous pressure fell considerably. Five of the six subjects developed signs of circulatory collapse 1–4 minutes after the end of the venesection: the blood pressure fell markedly, the heart rate was lowered to 36–50 beats per minute and the skin colour became ashen grey. The subjects felt very weak and vision was blurred; they complained of nausea and retching, sweated a good deal, and responded slowly to words of command; one subject lost consciousness. In two of these subjects who were untreated the heart rate remained slow and the blood pressure stayed below 100 mm. Hg for 30 minutes. During the next 12 hours their circulation was adequate in the recumbent position, though pallor persisted, but if they stood up they fainted. After 24 hours they could get out of bed safely.

Fig. 46 enables the sequence of events in hæmorrhage to be followed.

(1) *Blood Volume and Cardiac Output.*—The loss of blood, of course, decreases the *blood volume*; the venous pressure and the right auricular pressure fall (*e.g.* from 12 to 7·5 cm. saline); the venous return is diminished and the *cardiac output* is reduced (*e.g.* from 8·0 to 5·0 litres per minute).

(2) *Blood Pressure and Peripheral Circulation.*—The blood pressure varies as the product of the cardiac output and the peripheral resistance (p. 302); if all other factors remain unaltered halving the cardiac outp ut should halve the blood pressure. In fact, however, the blood pressure only falls by some 10 mm. Hg; this indicates that the peripheral resistance has undergone a compensatory increase of about 50%. This *vasoconstriction* is due partly to diminution of the normal tonic inhibitory activity of the sino-aortic nerves on the vasomotor and adrenaline-secreting centres.[1]

[1] In addition *pressor* afferent impulses pass up from the *chemoreceptors* which are stimulated by anoxia (p. 746). [Section of the vagi after a severe hæmorrhage may therefore cause a marked *fall* of pressure (Fig. 188).]

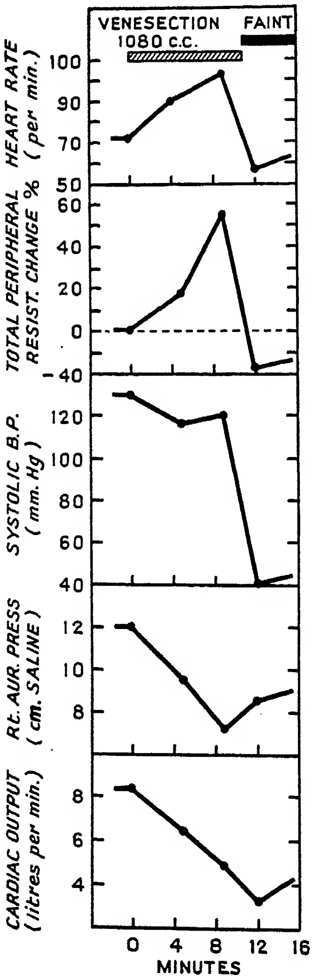


FIG. 46.—Circulatory Changes in severe Hæmorrhage in Man, followed by a Faint (cf. Figs. 158, 159, p. 275).

Records from above downwards are: heart rate per minute; percentage change in total peripheral resistance (0 on chart=initial level in subject); systolic blood pressure in mm. Hg; right auricular pressure in cm. of saline; cardiac output in litres per minute.

Rapid venesection to 1080 c.c. in 10 minutes was followed by a faint at the point indicated on chart. Note that the onset of the faint is associated with a marked fall of blood pressure (to 40 mm. Hg), slowing of the heart, and a great decrease in the total peripheral resistance; the last change is due to sudden marked dilatation of the muscle blood vessels. (Barcroft, Edholm, M'Michael and Sharpey-Schafer, *Lancet*, 1944.)

The *vasomotor centre* discharges more actively, leading to greater constriction of the skin vessels (accounting for the cold pale skin) and the splanchnic area, including the renal vessels ; the direct depressor effect of the hæmorrhage on the blood pressure is thus, for a time, largely kept in check. Adrenaline secretion is increased and acts peripherally to produce the same effects.

(3) *Blood Stores.*—The spleen contracts in some species (but probably not in man), discharging red cells stored in its pulp into the circulation. The pulmonary vessels constrict, expelling 150–200 c.c. of blood into the systemic circulation (p. 80).

By means of the compensatory reactions described (chiefly those in (2)) the blood flow to vital organs like the brain and heart is kept up to the best possible extent under the circumstances.

(4) *Heart Rate.*—The heart rate in hæmorrhage quickens (*e.g.* from 70 to 100 beats per minute) in spite of the associated fall of venous pressure which tends reflexly to slow the heart (p. 272) ; the cardiac acceleration is mainly brought about reflexly, via the sino-aortic nerves, by the fall of arterial pressure. Normal vagus tone is the result of reflex stimulation of the vagus centre by afferent impulses in the sino-aortic nerves set up by the arterial pressure ; the fall of blood pressure diminishes the afferent impulse stream, diminishes the stimulation of the vagus centre, decreases vagal "restraint" of the heart and so produces cardiac acceleration. The cardiac quickening has no beneficial effect on the cardiac output per minute as the heart cannot pump out more blood than it receives ; the output per *beat* falls from 110 c.c. to 50 c.c.

(5) *Respiration* is reflexly stimulated by the fall of blood pressure and the anoxia ; it becomes rapid, shallow or "sighing" in character. Because of the loss of circulating hæmoglobin, hæmorrhage leads to a decrease in oxygen-carrying power ; but this failure of oxygen transport is less serious than the circulatory failure described above. Most of the symptoms of hæmorrhage are rapidly overcome by an adequate transfusion of *plasma* (which does not increase oxygen-carrying power). It may also be pointed out that the far greater loss of oxygen-carrying power which occurs in severe anæmia (*e.g.* pernicious anæmia) is not associated with such grave symptoms.

(6) *Renal Changes.*—These are well illustrated by the following example. In a severe case of duodenal hæmorrhage in which the blood and plasma volume were reduced respectively to 4·1 and 2·9 litres, the glomerular filtration (inulin clearance) was only 25 c.c., and the renal plasma flow (diodrast clearance) 70 c.c., per minute. The decrease in renal plasma flow is due to the decreased cardiac output and lowered arterial blood pressure combined with marked renal vasoconstriction ; the filtration fraction in the glomeruli in the case just recorded was about 0·5 (normal 0·2), indicating selective vasoconstriction in the *efferent* glomerular vessels (cf. p. 25). It should be remembered that in hæmorrhage there may be constriction or closure of the interlobular arteries preventing the blood reaching the cortical glomeruli and tubules (p. 26), with diversion of the blood flow through the medulla ; the diodrast clearance may thus give an underestimate of the true renal blood flow but it measures the *effective* renal blood flow well enough (cf. p. 39). Seventeen hours after receiving a blood transfusion of 850 c.c., the blood and plasma volumes were raised in the patient to 5·4 and 3·8 litres. The glomerular

filtrate was 55 c.c. and the renal plasma flow 170 c.c. After complete recovery these latter values were 90 and 810 c.c. respectively.

The impaired renal circulation resulting from a hæmorrhage leads to a great decrease in urine flow and to nitrogenous retention in the blood. The level of blood urea 1 hour after the hæmorrhage was about 100 mg–% ; it fell gradually to normal as the renal blood flow rose.

(7) *Mechanism of the Faint.*—When the blood loss is very large the cardiac output falls further, but the outstanding event is the failure of the peripheral resistance. The exact cause is not known, but there is marked dilatation of the vessels in the skeletal muscles, leading to a collapse of the blood pressure (*e.g.* to 40 mm. Hg). There is also great cardiac slowing that is as yet unaccounted for. The diminished blood flow to the brain results in impairment or loss of consciousness (fainting). [Cf. *hæmorrhagic shock*, p. 341.]

Restoration of the Blood after Hæmorrhage.—The Table below and Fig. 47 summarize the changes in the blood during the natural recovery

Blood Constituent.	Normal Initial Values.	Remove 1070 c.c. of blood.	After 1 hr.	After 24 hrs.	After 72 hrs.
Plasma volume (c.c.)	2980	2430	2660	2900	3360
Red cell volume (c.c.)	2770	2250	2340	2360	2120
Total blood volume (c.c.)	5750	4680	5000	5260	5480
Plasma protein (g–%)	6·8	6·8	6·3	6·4	6·2
Total plasma protein (g.)	203	163	167	186	205
Added plasma protein (g.)	—	—	4	23	42
Hæmoglobin (g–%)	16·2	16·2	15·8	13·2	12·2

process after hæmorrhage. First the water, then the plasma proteins, and lastly the lost hæmoglobin and red blood cells are restored.

(1) *Fluid* begins to flow into the blood from the tissue spaces almost before the hæmorrhage is over ; after one hour there is some improvement in the plasma volume with a corresponding fall (owing to dilution) of the plasma protein and hæmoglobin concentrations. Plasma volume may be fully restored in 24 hours and may then rise to higher values to compensate for the volume of lost red cells which have not yet begun to regenerate. The mechanism of restoration of plasma volume is as follows: hæmorrhage results as stated in a fall of venous pressure (owing to underfilling of the vascular bed by the reduced blood volume) and consequently the capillary blood pressure falls. The balance between the capillary blood pressure and plasma protein osmotic pressure is disturbed (p. 18) with the result that fluid flows into the vessels from the tissue spaces.

(2) Additional *protein* begins to appear in the plasma within a few hours, and the lost plasma protein may be replaced within a few days (p. 137). The restoration of plasma protein may be due to rapid mobilization of reserve stores in depots (like the liver) and probably also to fresh plasma protein manufacture (in the liver). The return of protein to the plasma keeps up its colloid osmotic pressure and so enables the tissue fluid which has entered the circulation to be retained in spite of rising capillary blood pressure.

(3) The rise in *red cell* volume is delayed for some days, and then sets in as increased activity takes place in the red bone marrow; the effective stimulus to the marrow is anoxia resulting from the decreased oxygen-carrying power of the blood. As the normal blood condition is approached the anoxic stimulus becomes progressively weaker and the process of regeneration correspondingly slows down. The red cells and hæmoglobin are restored in 3–6 weeks if the diet is satisfactory and contains adequate amounts of the necessary raw materials (*e.g.* iron, protein) (p. 194). The total blood volume reaches normal value when the corpuscular regeneration is complete.

It is unlikely that contraction of the spleen in man plays any important part in replacing the lost circulating cells.

It should be stressed again that *immediately* after a hæmorrhage the total hæmoglobin content of the blood is reduced, but the hæmoglobin *percentage* is unchanged. As the plasma volume is restored, the hæmoglobin percentage

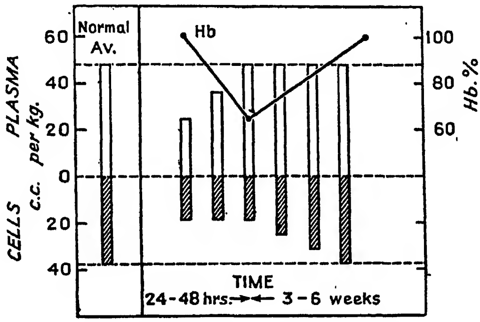


FIG. 47.—Diagram showing Recovery of Plasma and Cell Volume and Hæmoglobin after Hæmorrhage in Man. (Wright *et al.*, *Lancet*, 1938.)

The clear columns represent plasma volume and the hatched columns cell volume (in c.c. per kg.). The thick line (Hb) represents hæmoglobin percentage. Half the blood was lost. Note the rapid restoration of the plasma volume and resulting fall in hæmoglobin percentage. Later there is a slow recovery of the cell volume (and of total blood volume) and of percentage hæmoglobin.

correspondingly falls owing to simple dilution. As the red cells are regenerated the hæmoglobin percentage correspondingly rises (Fig. 47).

If the oxygen lack is very severe, as after a large hæmorrhage, the red marrow may itself be so adversely affected that the process of red cell restoration may be hampered.

Nucleated red cells may appear in the circulating blood during recovery, as well as immature cells (reticulocytes) or badly shaped forms (poikilocytes). Many small red corpuscles are found, tending to make the average diameter of the red cells lower than normal. The cells may be deficient in hæmoglobin (Fig. 94, A).

AIDS TO RECOVERY.—The full restoration of the plasma volume is hampered if the patient is dehydrated and the volume of his interstitial fluid is reduced; conversely, it is greatly facilitated if plenty of fluid is available from the tissue spaces. In cases of emergency, fluid must be introduced directly into the blood vessels so as to maintain an effective circulation.

(1) *Saline* injected intravenously, though it temporarily replaces the lost fluid, is rapidly eliminated, because it dilutes the proteins of the plasma; it

also lowers the viscosity of the blood. For these reasons it does not maintain the blood pressure for more than a short time and is useful only as a temporary expedient in an emergency (Fig. 48). Saline absorbed from the subcutaneous tissues or the bowel is very useful and is better retained, but the effects take longer to develop (cf. p. 62).

(2) *Plasma* injected intravenously is retained for a much longer time owing to the osmotic pressure of the contained plasma proteins; it raises both the plasma volume and the blood pressure and if given in adequate volume it restores the clinical condition practically to normal although it provides no oxygen-carrying power (Fig. 48). This observation shows that the grave effects of hæmorrhage are mainly due to *decreased circulating volume.*

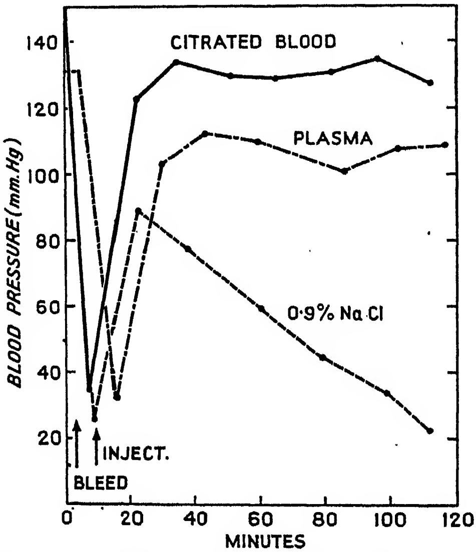


FIG. 48.—Effect of Transfusion on low Blood Pressure following Hæmorrhage.

Experiments on Cats. Standard severe initial hæmorrhage to lower blood pressure to 30–40 mm. Hg. Inject equal volumes (in different experiments) of citrated whole blood, citrated plasma, saline. Note varying degrees of recovery of the blood pressure. (Schweitzer *et al.*, *Lancet*, 1940, ii, 507.)

(3) The ideal method is *blood transfusion,* which provides a fluid of the right viscosity and osmotic pressure, and restores the blood volume, blood pressure and the *oxygen-carrying capacity* of the blood to their normal levels. Two essential precautions must be observed: the blood must be kept fluid, *i.e.* coagulation must be prevented (p. 143); and the compatibility of the offered red blood corpuscles must first be determined (p. 181).

Fig. 48 shows the relative effectiveness of equal volumes of 0·9% saline, plasma, and citrated blood administered immediately after a severe experimental hæmorrhage. Plasma (and serum) are only a little less satisfactory than blood in restoring the blood pressure permanently; the improvement with saline is very transient and the animals rapidly succumb.

Blood Volume in Disease.[1]—(1) In *nephritis* the total volume may be decreased if the plasma protein is lowered (decreased plasma volume) or if

[1] Gibson, *Ann. int. Med.*, 1940–41, *14*, 204.

the cell count is reduced (decreased cell volume). In the terminal uræmic stage of renal disease, marked fluctuations in plasma volume occur owing to loss of fluid from vomiting, diarrhœa, or hæmorrhage, and from the effects of the various therapeutic measures which are so energetically carried out.

(2) In *congestive heart failure* the total blood volume may increase by 20–50% in proportion to the severity of the clinical state; both plasma and cell volume increase, the latter sometimes to a greater extent. As the condition improves the total volume correspondingly declines (by 500–1500 c.c.), both plasma and cell volume participating in the change.

(3) In *diabetes mellitus* with marked glycosuria, the plasma volume is decreased, probably from loss of fluid as a result of diuresis. This anhydræmia contributes to the circulatory failure which may develop (cf. p. 925).

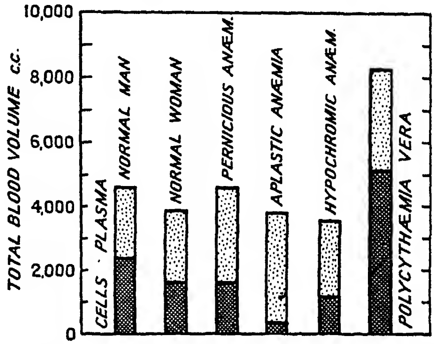


FIG. 49.—Plasma Volume, Cell Volume, and Total Blood Volume in various Disorders.

The normal values here shown are somewhat lower than those set out on p. 80. (Redrawn from Haden, *Bull. N.Y. Acad. Med.*, 1939, *15*, 291.)

(4) In some forms of *traumatic shock* there is marked loss of fluid into the tissue spaces or of blood in the injured areas (p. 342); great fluid loss may also occur in *burns* (p. 344) and in *high intestinal obstruction* (p. 108). In all these conditions plasma volume may be reduced. The plasma volume is also diminished in *adrenal insufficiency* (p. 955).

(5) In the *anœmias* the plasma volume may be increased as though to compensate for the great decrease in cell volume; this occurs in severe *pernicious anœmia* and in *aplastic anœmia*. The total blood volume may or may not be diminished. In *polycythœmia* the blood volume is regularly high because of the excessive volume of red cells; the plasma volume is unchanged (Fig. 49).

REGULATION OF H+ ION CONCENTRATION [ACID-BASE BALANCE] OF BLOOD[1]

Pure water consists almost entirely of undissociated molecules of H_2O; but very slight dissociation does occur giving rise to hydrogen ions (H^+) and hydroxyl ions (OH').[2]

$$H_2O \rightleftharpoons H^+ + OH'.$$

An H^+ ion is an H atom minus one (negatively charged) electron; an OH' ion is an OH particle which has gained one electron. The reaction of any

[1] Haldane and Priestley, *Respiration*, new edn., Oxford, 1935. Gamble, *Extracellular Fluid*, Harvard, 1949. Singer and Hastings, *Medicine*, 1948, *27*, 223.

[2] According to the teachings of modern physical chemistry, hydrogen ions do not exist as H^+ but as H_3O^+ (*i.e.* as $H^+ + H_2O$). Fortunately the physical chemists are agreeable to the ion being treated as H^+ in physiological discussions.

fluid (*i.e.* its acidity or alkalinity) depends on the relative *number* (not weight) of H^+ ions and OH' ions; if there are more H^+ ions the solution is acid, if there are more OH' ions the solution is alkaline. Pure water is a neutral solution: it therefore contains equal *numbers* of H^+ and OH' ions. The *concentration of H^+ (cH)* in water is expressed as gramme equivalents of ionic H per litre of pure water at 22° C.; it is written when so expressed as $[H^+]$; the value is $10^{-7}=\frac{1}{10^7}=\frac{1}{10\text{m.}}$ (where m. stands for million). The equivalent weight of H is one. Therefore $[H^+]=10^{-7}$ means that there is 1 g. of H in the ionic form in 10 m. L of water (or $\frac{1}{10\text{m.}}$ g. in 1 L of water).

The concentration of OH' in water is also 1 g. equivalent of ionic OH in 10 m. L, *i.e.* $[OH']=10^{-7}$. The equivalent weight of OH is 17; therefore the *weight* of OH' in water is 17 g. in 10 m. L of water. But though the *weight* of OH' in water is 17 times as great as the weight of H^+ the *number* of each present is the same; the use of the gramme-equivalent method of notation enables us to compare the relative *number* of H^+ and OH' while avoiding the complication of their difference in weight. The gramme-equivalent notation likewise enables us to compare the *number* of chemically combining units (p. 12).

In all aqueous solutions whatever their reaction the product of $[H^+]$ and $[OH']$ is a constant (its value is 10^{-14}). In the case of pure water:

$$[H^+]\times[OH']=10^{-7}\times 10^{-7}=10^{-14}$$
$$=\tfrac{1}{10\text{m.}}\times\tfrac{1}{10\text{m.}}=\tfrac{1}{100\text{m.m.}}\text{ (see footnote 1).}$$

It follows that if we know the H^+ ion concentration of any aqueous solution, we can calculate the OH' ion concentration, because the product of the two must be 10^{-14}. We can then compare the H^+ ion and OH' ion concentration, see which is greater and so know the reaction of the fluid.

$\therefore$ $[H^+]=10^{-7}$; $[OH']=10^{-7}$=neutral solution.
$[H^+]=10^{-6}$; $[OH']=10^{-8}$=acid solution; $[H^+]$ is 100 times $[OH']$.
$[H^+]=10^{-8}$; $[OH']=10^{-6}$=alkaline solution; $[OH']$ is 100 times $[H^+]$.

Note that *lowering* the size of the index *increases* the concentration, 10^{-6} ($\frac{1}{1\text{m.}}$) being a 100 times greater concentration than 10^{-8} ($\frac{1}{100\text{m.}}$).

A normal (N.) solution [2] of any acid contains 1 g. of displaceable hydrogen per litre. A strong acid dissociates almost completely, so that in the case of HCl, for example, a normal solution would contain nearly all its 1 g. of replaceable H per L in the ionized form, *i.e.* $[H^+]$ is approximately 1. $[OH']$ would, of course, be 10^{-14}. A decinormal (0·1 N.) solution of HCl would contain nearly all its 0·1 g. of replaceable H per L in the ionic form, *i.e.* $[H^+]$ is $0{\cdot}1=10^{-1}$, and $[OH']=10^{-13}$. Similarly, a normal solution of any alkali contains 1 g. equivalent of OH (17 g. OH) per litre. A strong alkali dissociates almost completely, so that in the case of N.NaOH the solution contains (nearly) all its 1 g. equivalent of OH per L in the ionized form; $[OH']=1$, and

[1] Note again that m.=million, and m.m.=million million.

[2] A *normal* (N.) solution of any acid or alkali contains the equivalent weight of the substance in 1 L of water. Thus N.HCl contains 36·5 g. per L; N.H_2SO_4 contains $\frac{\text{mol. wt.}}{2}=\frac{98}{2}$=49 g. per L; N.NaOH contains 40 g. per L.

$[H^+]=10^{-14}$. Decinormal (0·1 N.) NaOH contains its 0·1 g. equivalent of replaceable OH in the ionic form, *i.e.* $[OH']=0{\cdot}1=10^{-1}$; $[H^+]=10^{-13}$.

In the case of a *weak* acid (or alkali) ionization is very *incomplete*, and only a small proportion of the replaceable H (or OH) of the compound is in the ionic form. This is illustrated in the Table below.

DEGREE OF DISSOCIATION OF ACIDS AND BASES.

Acids.	Percentage Dissociation.	Bases.	Percentage Dissociation.
0·1 N. HCl . .	91·0	0·1N. NaOH	91·0
0·1 N. Acetic .	1·34	0·1N. NH_4OH	0·4
0.1 N. Carbonic .	0·17		

Note that decinormal HCl or NaOH dissociates almost completely (91%). A weak acid like carbonic in 0·1 N. solution is only dissociated to the extent of 0·17%. Thus, though decinormal solutions of HCl and H_2CO_3 have the same amount of *displaceable* H per litre, they differ enormously in their *ionic* hydrogen concentration. Thus the $[H^+]$ of 0·1 N. HCl compared with that of 0·1 N. H_2CO_3 is $\frac{91}{0{\cdot}17}$, *i.e.* it is 535 times as great. Similarly the $[OH']$ of 0·1 N. NaOH compared with that of 0·1 N. NH_4OH is $\frac{91}{0{\cdot}4}$ or 228 times as great.

To avoid the negative indices of the above method, Sørensen has introduced another nomenclature, that of pH; the letter p signifies that the negative exponent to the base 10 is employed (*i.e.* $p\text{H}=-\log_{10}c\text{H}$). Thus :

N. HCl	$[H^+]=1$	$=10^0$	pH=0
0·1 N. HCl	„ $=0{\cdot}1$	$=10^{-1}$	„ =1
0·01 N. HCl	„ $=0{\cdot}01$	$=10^{-2}$	„ =2
Pure water	„	$=10^{-7}$	„ =7
0·1 N. NaOH	„	$=10^{-13}$	„ =13
N. NaOH	„	$=10^{-14}$	„ =14

A solution of pH 7 is *neutral*; if the pH exceeds 7, the solution is *alkaline*; if the pH is less than 7, the solution is *acid*. Most people, however, find it difficult to think in terms of negative indices; such people should always translate pH values into H^+ ion concentration. Every rise of 1 in pH means lowering the concentration of H^+ ions to *one-tenth* its previous value. Fig. 50 shows that a rise of 0·3 in the pH value reduces the H^+ ion concentration to *half*; a rise of 0·6 in pH reduces the H^+ ion concentration to *one-quarter*; a rise of 0·9 in pH reduces the H^+ ion concentration to *one-eighth*.

NORMAL BLOOD REACTION.—The normal pH of blood (measured electrometrically) is 7·35–7·45 (mean 7·4) at 37° C., *i.e.* slightly on the alkaline side of neutrality. The limits of pH compatible with survival are pH 7 on the acid side and pH 7·8 on the alkaline side. By *acidæmia* (*acidosis*) is meant increased H^+ ion concentration of blood beyond the normal limits. By

alkalæmia (*alkalosis*) is meant a diminished H^+ ion concentration below the normal limits.

In this book the terms acidæmia or alkalæmia (without qualification) are, as a rule, only used when the change is *demonstrable by physical measurement* ; otherwise the phrase " *tendency* to acidæmia or alkalæmia " is used.

Ketosis signifies the presence in excess of certain fatty acids :

β-hydroxy-butyric acid, $CH_3CH(OH)CH_2COOH$
Acetoacetic acid, CH_3COCH_2COOH

in the blood, irrespective of the blood reaction (cf. p. 869).

The *alkali reserve of the plasma* is the bicarbonate content of the plasma at the CO_2 tension present in the arterial plasma (which is the same as in the

*p*H	cH expressed as g. of H^+ ions per litre.
7·0	$\frac{1}{10 \text{ million}}$
7·3	$\frac{1}{20 \text{ million}}$
7·6	$\frac{1}{40 \text{ million}}$
7·9	$\frac{1}{80 \text{ million}}$
8·0	$\frac{1}{100 \text{ million}}$

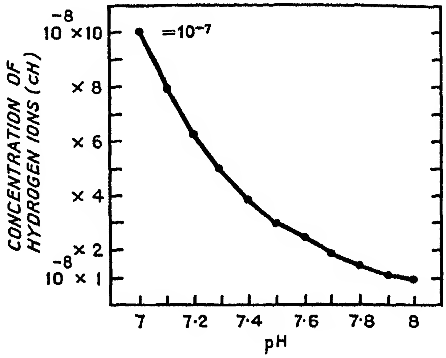


FIG. 50.—Relationship of Concentration of Hydrogen Ions (cH or [H^+]) to *p*H. (Graham and Morris, *Acidosis and Alkalosis*, 1933.)

alveolar air ; normal CO_2 tension=40 mm. Hg). The plasma is equilibrated at the appropriate CO_2 pressure and then treated with acid and exposed to a vacuum in the Van Slyke apparatus (cf. p. 416). Suppose that 100 c.c. of plasma give off 63 c.c. of CO_2. The amount of CO_2 that was present in *solution* is calculated from a knowledge of the CO_2 pressure and solubility of CO_2 in plasma ; the normal value is 3 c.c. per 100 c.c. Therefore 63—3=60 c.c. of CO_2 were present in the plasma in the *combined* form, which for all practical purposes means as *bicarbonate*. Note that the plasma bicarbonate (per 100 c.c.) is expressed not as a weight, but as the *volume* of CO_2 which is evolved from it.

Volumes of CO_2 (per 100 c.c.) are converted into milliequivalents (m.Eq.) per L by dividing by 2·22. Thus the normal values for plasma are :

Dissolved CO_2=3 c.c. per 100 c.c.
=1·35 m.Eq. per L (as $H.HCO_3$).
Combined CO_2=60 c.c. per 100 c.c.
=27 m.Eq. per L (as $B.HCO_3$)
(B=Base=almost entirely Na^+) [1]

[1] In the text bicarbonate is written as $BHCO_3$ or $B.HCO_3$, carbonic acid as H_2CO_3 or $H.HCO_3$, as seems most appropriate.

Buffer substances are those which hinder changes in the reaction of a fluid. The word is derived from a German word meaning tampon, a buffer substance being regarded as " mopping up " H^+ or OH' ions and thus preventing them from accumulating. The mode of action of buffers will be made clear later (p. 92).

The chief *acid* present as such in blood is the CO_2 which is present in solution giving rise to H_2CO_3 which yields H^+ ions :

$$H_2O + CO_2 \leftrightharpoons H.HCO_3 \leftrightharpoons H^+ + HCO_3'.$$

In addition, plasma protein and hæmoglobin exist to some extent in blood in the acid form, but their degree of ionization is negligible.

The only *alkali* present as such in blood is the bicarbonate ($BHCO_3$). Let us assume that the base is Na^+, *i.e.* for B substitute Na^+. The dissociation occurs thus :

$$\text{(i)}\quad NaHCO_3 \rightleftharpoons Na^+ + HCO_3'.$$

The HCO_3' then " mops up " H^+ ions (derived from dissociation of water) to form the weakly dissociating acid, H_2CO_3 :

$$\text{(ii)}\quad HCO_3' + H^+ \rightleftharpoons H_2CO_3.$$

Reaction (ii) causes further dissociation of water to yield *an excess of OH′ ions* ; the total reaction can be summarized thus :

$$\text{(iii)}\quad Na^+ + HCO_3' + H^+ + OH' \rightleftharpoons Na^+ + H_2CO_3 + \mathbf{OH'}.$$

The essential point to remember is that the presence of bicarbonate increases the OH' concentration of the solution, *i.e.* increases its alkalinity.

As the reaction of blood depends on the relative numbers of H^+ and OH', and as the principal source of H^+ is H_2CO_3 and the principal source of OH' is bicarbonate ($BHCO_3$), then the H^+ ion concentration varies as the ratio of H_2CO_3 to $BHCO_3$ (cf. p. 92).

Regulation of Blood Reaction.—Many factors are normally operating to disturb the blood reaction from its normal range ; the compensatory mechanisms are however so efficient that in health detectable changes in blood reaction occur rarely.

Acids are constantly being formed in the body as a result of metabolic activity : CO_2 results from the oxidation of all foodstuffs ; H_3PO_4 and H_2SO_4 from the oxidation of the P and S of ingested protein. Lactic acid is formed in the muscles during activity, but normally it is disposed of as rapidly as it appears. In very strenuous exercise, lactic acid is formed in greater amounts than the muscles can deal with, and consequently overflows into the blood. Under abnormal conditions β-hydroxybutyric acid, acetoacetic acid, and possibly other acids not hitherto recognized may appear in the circulation.

Basic radicals like Na, K, Ca, Mg are taken in large amounts in an ordinary diet, especially in vegetable food, and they must be disposed of to prevent excessive alkalinity of the blood.

The methods available in the body for maintaining blood reaction are of two kinds : (i) the *physico-chemical* processes occurring in the blood ; (ii) the *vital* reactions of the breathing and the kidneys.

I. Physico-Chemical Reactions.—These depend on : (i) the buffering action of the plasma $BHCO_3$; (ii) methods of transporting CO_2 in forms other than H_2CO_3 (*i.e.* non-acid forms).

(1) Relation of Plasma Bicarbonate to pH.—As explained (p. 91), blood reaction depends on the ratio of H_2CO_3 to $BHCO_3$ (cf. Fig. 58, A). For plasma, the Henderson-Hasselbalch equation expresses this relationship quantitatively :

$$pH = 6{\cdot}1 + \log \frac{B.HCO_3}{H.HCO_3}.$$

$$\text{Normally } \frac{B.HCO_3}{H.HCO_3} = \frac{60}{3} \text{ vols.\%} = \frac{27}{1{\cdot}35} \text{ m.Eq./L} = 20.$$

$$\text{As } \log 20 = 1{\cdot}3 ;\ pH = 6{\cdot}1 + 1{\cdot}3 = 7{\cdot}4.$$

The following deductions may be drawn from the equation :

(i) As long as $\frac{B.HCO_3}{H.HCO_3}$ remains constant, changes in the absolute values do not alter pH. Thus, if both values are doubled, *e.g.* $B.HCO_3=120$, $H.HCO_3=6$ vols.%, or, if both values are halved, *e.g.* $B.HCO_3=30$, $H.HCO_3=1{\cdot}5$ vols.%, the ratio remains 20 (log=1·3) and the pH remains 7·4.

(ii) *Diluting* the plasma does not alter its pH.

(iii) Any change in *either* $B.HCO_3$ *or* H_2CO_3 alters pH : thus increase of $B.HCO_3$ or decrease of $H.HCO_3$ produces alkalæmia ; decrease of $B.HCO_3$ or increase of $H.HCO_3$ produce acidæmia.

(iv) *Unequal* changes in $B.HCO_3$ or $H.HCO_3$ alter pH.

(v) Changes in $B.HCO_3$ and $H.HCO_3$ in *opposite* directions markedly change pH.

One example may be given : suppose that plasma $B.HCO_3$ is reduced to 27, and $H.HCO_3$ remains unchanged at 3 vols.%.

$$\frac{B.HCO_3}{H.HCO_3} = \frac{27}{3} = 9 ;\ \log 9 = 0{\cdot}95$$

$$pH = 6{\cdot}1 + 0{\cdot}95 = 7{\cdot}05.$$

(2) Buffering Action of Plasma Bicarbonate.—When any non-volatile acid like lactic, phosphoric, or sulphuric acid enters the blood stream it at once interacts with the bicarbonate present in the plasma as follows, *e.g.*:

$$\underset{\text{(Lactic acid.)}}{H^+ + L'} + Na.HCO_3 \rightarrow \underset{\text{(Sodium lactate.)}}{Na^+ + L'} + H.HCO\ .$$

In the case of lactic acid, sodium lactate and carbonic acid are formed. The advantage of this reaction to the body is clear. Lactic acid is a comparatively strong acid, *i.e.* an acid which dissociates freely and gives rise to many H^+ ions, and thus makes any solution in which it is present considerably more acid ; it is also a non-volatile acid, *i.e.* one which is excreted slowly in the urine. By the interaction with sodium bicarbonate, lactic acid is converted into a neutral salt (sodium lactate) and H_2CO_3 is liberated. Now H_2CO_3 is a much weaker acid than lactic, giving rise to correspondingly fewer H^+ ions ; furthermore, it is a volatile acid which is readily eliminated in the lungs. Sodium bicarbonate has therefore acted as a *buffer* substance ; it has cause H^+ ions to be " mopped up," fewer are left in the solution, and the reaction of the blood has altered to a much slighter extent than otherwise would have been the case. That is why the term " alkali reserve " was introduced to describe the bicarbonate of the plasma (cf. p. 90).

It follows that as long as $NaHCO_3$ is present in the plasma, no acid stronger than carbonic can exist in the blood as such. The above type of buffering reaction takes place during very violent *muscular exercise* (p. 439), in severe *diabetes mellitus* (p. 924), and in *starvation* (p. 901).

This reaction alone, without further support, may still permit of considerable *p*H changes because as a result $BHCO_3$ is reduced and H_2CO_3 is simultaneously increased. To take an extreme case: suppose half the $BHCO_3$ is used up liberating 30 vols. of CO_2 (as H_2CO_3), then

$$\frac{B.HCO_3}{H.HCO_3}=\frac{30}{3+30}=\text{approx. } 1. \quad \text{Log } 1=0;$$

$$\therefore\ p\text{H}=6{\cdot}1+0=6{\cdot}1 \text{ (a fatal degree of acidæmia) (cf. Fig. 58, B).}$$

In the body the CO_2 evolved is immediately *mopped up in the blood* by the reactions mentioned below (it is chiefly taken up by hæmoglobin) and it is *finally eliminated in the lungs* (cf. Fig. 58, C, D).

(3) Buffering Action of Blood Proteins (mainly Hæmoglobin).—The acid which is formed in largest amounts in the body is H_2CO_3 itself, derived (i) principally as an end-product of normal metabolism (200 c.c. of CO_3 per minute at rest, up to 4 L in maximal exercise); (ii) from the buffering reaction described above.

When more CO_2 is formed, the CO_2 tension in the blood rises, more CO_2 is dissolved, more H_2CO_3 is formed, and more H^+ ions are liberated. If all the CO_2 had to be carried from the tissues to the lungs in solution, the requisite CO_2 tension would be great, and the rise of H^+ ion concentration serious. Most of the CO_2 is however removed from solution (and is then no longer present as H_2CO_3). The reactions concerned in this process are fully described on pp. 414 *et seq*., but may be summarized thus:

(i) CO_2 combines with the amino groupings of hæmoglobin to form carbamino-hæmoglobin.

(ii) CO_2 combines with the base (K) of hæmoglobin to form the *alkaline* substance $KHCO_3$.

Both these reactions are greatly facilitated by the simultaneous reduction of hæmoglobin.

Now hæmoglobin=hæm+globin. Hæm is an iron-porphyrin; depict it as Por : Fe^{++}. Globin can be depicted as $\left[R\begin{matrix}\diagup NH_2\\ \diagdown COOH\end{matrix}\right]$. Using these symbols the removal of excess CO_2 from solution occurs thus:

$$CO_2+\text{Por}:\ Fe^{++}O_2\left[R\begin{matrix}\diagup NH_2\\ \diagdown COOK\end{matrix}\right]\longrightarrow \text{Por}:\ Fe^{++}\left[R\begin{matrix}\diagup NH.COOH\\ \diagdown COOH\end{matrix}\right]+KHCO_3+O_2.$$

(Oxygenated potassium hæmoglobinate) (Reduced hæmoglobin combined with CO_2 in the carbamino form)

Certain subsidiary reactions (secondary buffering, chloride shift (p. 419)) help to preserve the reaction of the plasma.

By the conversion of CO_2 into bicarbonate, an acid (H_2CO_3) is converted into and transported as an *alkaline* substance; thus, when the concentration of H_2CO_3 is increased, $BHCO_3$ is simultaneously increased (though not to the same proportional extent), and so the $BHCO_3/H_2CO_3$ ratio which determines *p*H is relatively little altered. Direct measurement shows that the *p*H of mixed venous blood at rest is only 0·01–0·02 lower than that of arterial

Fig. 51 shows that the phosphate mechanism is equally effective in excreting acid and base.

Cellular Origin of Urinary Phosphate. Buffering Action of Cell Base.—These questions can be conveniently discussed here, as they are closely related to the regulation of the blood reaction. Some of the urinary phosphate is derived from the food; much of it, however, is derived from the body *cells.* In the case of the red cells, out of a total of 110 m.Eq./L of base, 30 m.Eq. are combined with *organic* phosphoric esters and only 1 m.Eq. is combined with *inorganic* phosphate. The large phosphate ester anions cannot penetrate the cell and are locked inside; the smaller inorganic phosphate ions

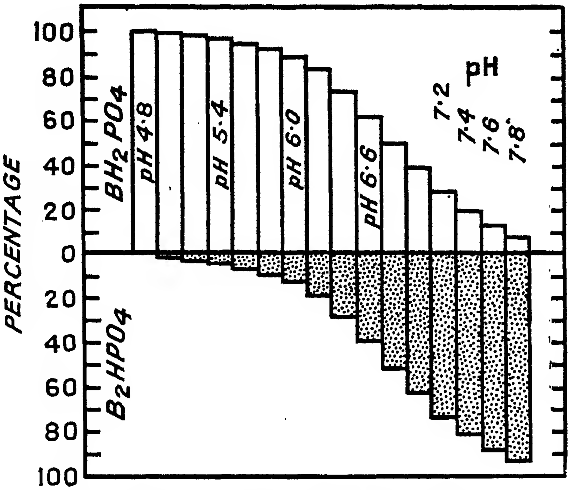


FIG. 51.—Relationship of Ratio of BH_2PO_4 to B_2HPO_4 to *p*H of Plasma and Urine.

Clear columns: BH_2PO_4. Stippled columns: B_2HPO_4.
The *p*H value at different ratios is indicated. Note values at normal *p*H of plasma (7·4). (After Gamble, *Extracellular Fluid*, Harvard University Press, 1949.)

can move in and out of the cells. The normal preponderance of the large organic over the small inorganic phosphate anions depends on a balance being struck between the *synthesis* of esters (phosphorylation) and their *breakdown* (phosphorolysis); the normal balance is greatly in favour of the synthesis of esters. In conditions of *acidæmia*, however, the equilibrium is disturbed and moves in favour of breakdown; the inorganic phosphate thus increases in the cells, diffuses into the plasma and can be excreted by the kidneys. Total red cell phosphate may fall, *e.g.* from 30 to 15 m.Eq./L. The phosphate in the cells at *p*H 7·2 is combined with *two* equivalents of base (potassium); when phosphate is excreted in a *urine of maximum acidity* it is excreted as BH_2PO_4, *i.e.* it is combined with only *one* equivalent of base (B). Consequently, for each equivalent of phosphate excreted in the urine there is one of potassium left behind in the plasma. This K^+ is available to combine with those acids which have caused the acidæmia; thus the

hydrolysis of organic cell phosphate, by causing K^+ to be released, helps to combat the condition. As all the body cells contain large amounts of phosphoric esters, it is reasonable to assume that similar contributions of base (K^+) are made available in acidosis by all the tissues. It is important to bear in mind that potassium which is freed from the cells may be excreted by the kidneys in combination with the normal acids; hence acidæmia may lead to a dangerous degree of potassium depletion of the tissues; this depletion may need to be made good by treatment.

In alkalæmia, the reverse of these reactions takes place: synthesis of cell esters is favoured, and total cell phosphate may rise, *e.g.* from 30 to 45 m.Eq./L.

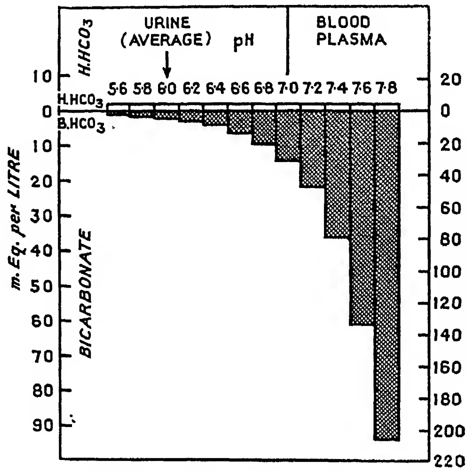


FIG. 52.—Relationship of Carbonic Acid–Bicarbonate Ratio to *p*H of Urine and Plasma.

H_2CO_3 (clear columns) above horizontal line; BHCO (hatched) below horizontal line.

Ordinate: *Left*, milliequivalents per litre (m.Eq./L). *Right*, volumes of CO_2 per cent. To convert m.Eq./L into vols. of CO_2 per cent., multiply by 2·2 (*e.g.* 20 m.Eq./L of $BHCO_3$=44 c.c. of CO_2%).

Figures above H_2CO_3 line=*p*H values. Note that average urinary *p*H is 6·0; normal plasma *p*H is 7·4. (After Gamble, *Extracellular Fluid*, Harvard University Press, 1949.)

Extra base is thus withdrawn from the plasma into the cells relieving the alkalæmia without altering the *p*H of the tissues. The kidneys need only excrete the base which is in excess of the amount used to combine with the increased phosphoric esters retained in the cells.

(ii) *Bicarbonate Mechanism.*—The bicarbonate mechanism is really effective only when compensating for states of alkalæmia (Fig. 52). At a urinary *p*H of 7·8, the maximal urinary $BHCO_3$ concentration is attained corresponding to over 90 m.Eq./L (=over 200 vols.% of CO_2 as $BHCO_3$); *i.e.* the $BHCO_3$ is concentrated in the urine to over three times its normal concentration in the plasma; very large amounts of alkali are thus excreted. In states of acidæmia the $BHCO_3$ excretion is negligible, *i.e.* it is in effect completely reabsorbed, no base therefore being eliminated.

(2) THE AMMONIUM MECHANISM.—The epithelium of the renal tubules manufactures NH_3 from *amino-acids*[1] which are present in the glomerular filtrate; the NH_3 is then passed into the lumen of the tubule to neutralize acid substances filtered out by the glomeruli, forming *ammonium* salts. A certain amount of fixed base is consequently freed and can be transferred from the lumen of the tubule back into the blood. For example: lactic acid is neutralized in the blood by $NaHCO_3$ (which is destroyed) to form Na lactate; in the tubules it reacts with NH_3 made by the kidney to form ammonium lactate; the Na^+ which is freed returns to the blood to re-form $NaHCO_3$ (Fig. 54). The ammonium mechanism thus spares the fixed base (and the alkali reserve) of the plasma.

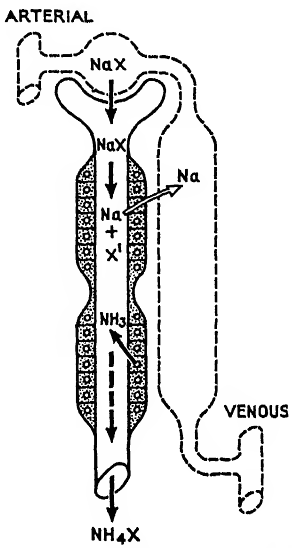


FIG. 53.—Diagram illustrating Ammonium Mechanism in Kidney (David Slome).

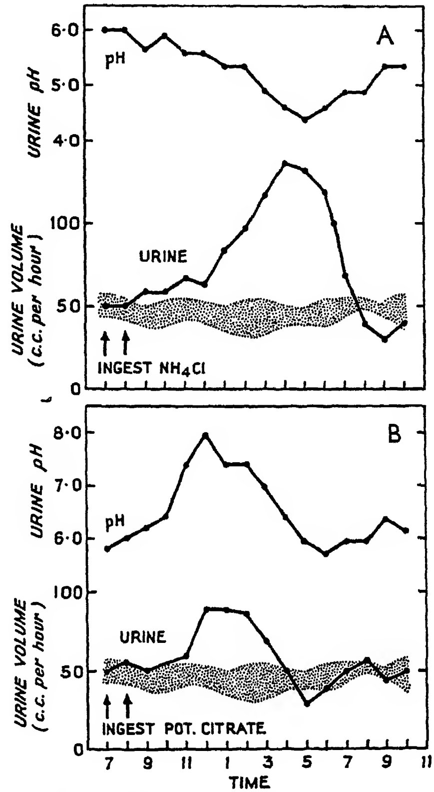


FIG. 54.—Regulation of Reaction of Urine in Man. (Experiments by John Marks.)

The shaded area in each figure represents the maximum and minimum output of urine during each hour of the day in a period of 28 days on a constant water intake. Time, from 7 a.m. to 11 p.m.
A. Upper Figure: at each arrow ingest 3 g. of ammonium chloride.
B. Lower Figure: At each arrow ingest 2 g. of potassium citrate.
There is an increase in urinary acidity (fall of pH from 6·0 to 4·5) in upper figure, and a decrease in urinary acidity (rise of pH from 6·0 to 8·0) in lower figure. (Citrate is combusted in the body to CO_2 and the retained base (K) is converted into bicarbonate ($KHCO_3$).) Diuresis occurs in both experiments.

The NH_3 mechanism has at its disposal not only all the amino-acids derived from the protein of the food, but when necessary the amino-acids

[1] It is claimed that the kidney forms NH_3 most readily from glutamine (the amide of glutamic acid); but as stated, the kidney can liberate NH_3 from amino-acids in general.

of tissue protein also. In so far as NH_3 of amino-acids is used by the renal tubules for neutralization purposes less is available for urea formation in the liver (p. 886); an increase in the excretion of ammonium salts in the urine reduces the amount of urea excreted to a corresponding extent.

In conditions of acidæmia the kidney makes more NH_3 to deal with the excess of acid radicals present; in alkalæmia, as ample base is already available, the kidney makes less NH_3 or none at all. The urinary "ammonia" (it is always present as ammonium salt) is thus increased in acidæmia and decreased in alkalæmia.

When the urine *p*H is less than 6·0, *organic acids* (*e.g.* acetoacetic acid, β-hydroxybutyric acid) can be excreted as such, *i.e.* their elimination does then not involve simultaneous loss of base.

Response of the Kidney to Acidæmia and Alkalæmia.—This is illustrated by the following experiments:

(i) *Ingest NH_4Cl.* — This salt is treated by the liver as though it were an amino-acid; it is deaminized, *i.e.* NH_3 is split off and HCl liberated. The NH_3 is converted in the usual way into urea; the HCl induces an acidæmia. The kidney responds by excreting a urine of maximal acidity (Fig. 54, A) with a progressively rising NH_4^+ content (Fig. 55).

(ii) *Ingest Potassium Citrate.*—The citrate radical is converted in the body to CO_2, and the freed K^+ combines with HCO_3' to form $KHCO_3$, producing an alkalæmia. The kidney responds by secreting a urine of maximal alkalinity (Fig. 54, B).

(iii) *Voluntary Overventilation.* — Alkalæmia is induced because the CO_2 tension in the alveolar air and consequently in the arterial blood is reduced, lowering the volume of CO_2 in solution and therefore the amount of H_2CO_3 and the number of H^+ ions (*i.e.* the ratio $H_2CO_3/BHCO_3$ is disturbed to the alkaline side). Fig. 56 shows how the kidney responds by excreting a highly alkaline urine; the NH_4^+ excretion is markedly cut down. Similar reactions occur when overventilation is produced by anoxia (p. 401), by raised body temperature (p. 478), in cardiac failure (p. 461), by lesions of the

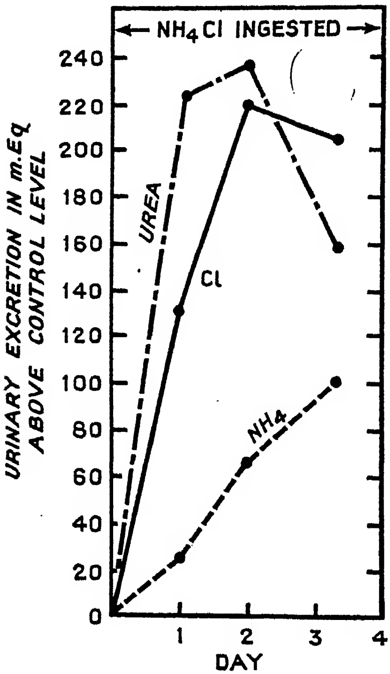


FIG. 55.—Effect of Ingestion of Ammonium Chloride on Urinary Excretion of Urea, Chloride, and Ammonium salts. (After Gamble, *Extracellular Fluid*, Harvard University Press, 1949.)

226 m.Eq. of Cl' as NH_4Cl were ingested daily. The ordinate indicates the *increase* in excretion of urea, Cl' and NH_4^+ in the urine above the control values.
The increased urea excretion is due to the NH_3 of the NH_4Cl being converted into urea. The released Cl' is excreted partly combined with fixed base and partly with extra NH_4^+ *made in the renal tubules*. Note the increased urinary excretion of NH_4^+.

central nervous system (in the hypothalamic region) (p. 717), or even by "simple" hysteria.

(iv) *Effect of Meals.*—Fluctuations in the reaction of the urine take place in relation to meals. Following the first meal the urine commonly becomes less acid (*alkaline tide*), and a few hours later it becomes more acid (*acid tide*). Some workers interpret the alkaline tide as a compensatory reaction to the alkalæmic tendency produced by the secretion (and loss from the blood) of the acid of the gastric juice; similarly they ascribe the acid tide to the acidæmic tendency which results from the secretion of the alkaline pancreatic juice, succus entericus, and bile (cf. p. 397 and Fig. 242). The urinary reaction, however, also seems to depend to a considerable extent on the *kind* of food taken at the meal, *i.e.* whether it is predominantly acid- or base-producing food.

Pathological Changes in Blood Reaction.—A deviation of the blood reaction to the acid side of the normal is called *acidæmia* or *acidosis*; a deviation to the alkaline side is called *alkalæmia* or *alkalosis*. Changes in blood reaction can also be classified as "*metabolic*" or "*respiratory.*"

(i) In "*metabolic*" changes in the blood reaction, there is a *primary change in the plasma bicarbonate.* In "metabolic" acidæmia there is an initial decrease in plasma bicarbonate; in "metabolic" alkalæmia there is an initial increase in plasma bicarbonate (Fig. 57, A, B, C, D).

(ii) In "*respiratory*" changes in blood reaction the *change in H_2CO_3 is primary.* In "respiratory" acidæmia there is an initial increase in H_2CO_3; in "respiratory" alkalæmia there is an initial decrease in H_2CO_3 (Fig. 57, E, F, G, H).

In all forms of alkalæmia and acidæmia compensation is "attempted" with varying degrees of success by appropriate changes in the blood buffers

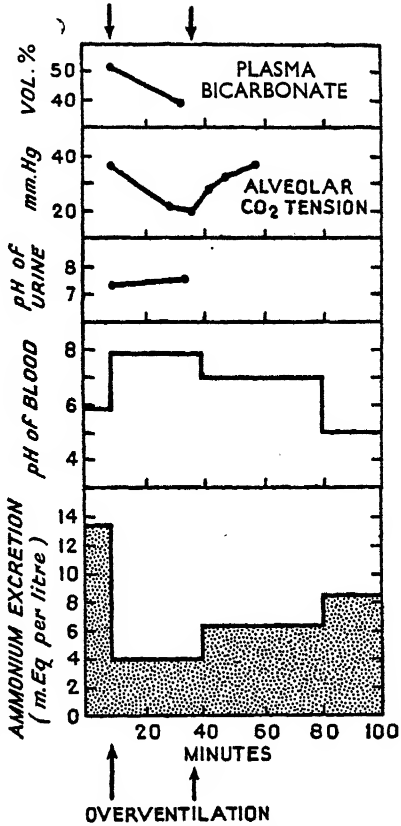


FIG. 56.—Effect of Prolonged Moderate Voluntary Hyperpnœa on Alveolar Air, Blood and Urine, showing compensatory reactions to Alkalæmia. (Grant and Goldman, *Amer. J. Physiol.*, 1920.)

Abscissa = Time in minutes.
Ordinates from above downwards:
Plasma bicarbonate expressed as volumes of CO_2 per 100 c.c.
Alveolar CO_2 tension in mm. Hg.
*p*H of urine } a rise in *p*H means increased
*p*H of blood } alkalinity.
Ammonium of urine in m.Eq./L.
Forced respiration begins at 10 mins. and ends at 38 mins. (during period indicated by the arrows). At 31 mins., tingling in foot; at 34 mins., Trousseau's sign present; at 35 mins., both hands in spontaneous spasm; at 36 mins., slight spasm of feet (see pp. 408, 1005); at 42 mins., temporary dizziness and trembling of hands on standing up after period of forced respiration is over.
Note that *p*H of blood rises in spite of compensatory reactions of kidney.

(especially hæmoglobin), the breathing, and the activity of the kidneys. The completeness of the compensation depends on the extent to which any or all of these mechanisms can be brought into play.

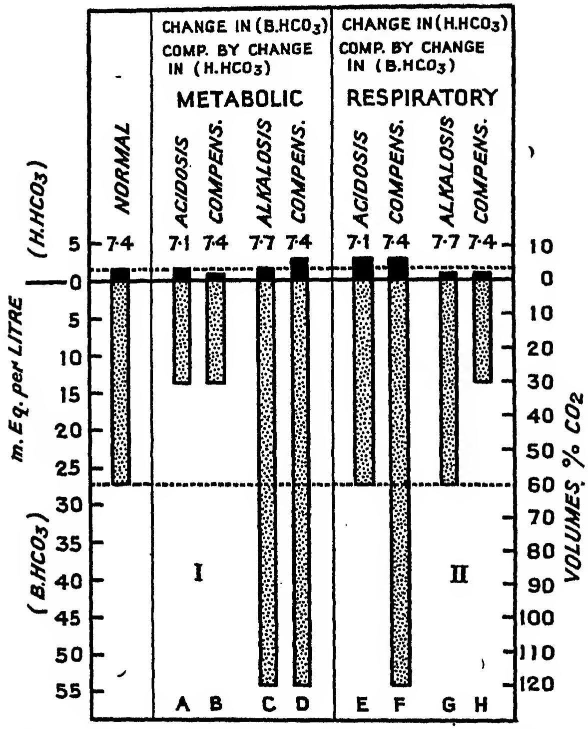


FIG. 57.—Relation of Carbonic Acid–Bicarbonate Ratio to *p*H of Plasma in Metabolic and Respiratory Acidosis and Alkalosis, compensated and uncompensated. (After Gamble, *Extracellular Fluid*, Harvard University Press, 1949.)

Note that in this Fig., H_2CO_3 and $BHCO_3$ are expressed as m.Eq./L; in Fig. 58 they are expressed as volumes of CO_2 per 100 c.c.

I. In "*metabolic*" changes of blood reaction the *change in $BHCO_3$ is primary*.

A. Uncompensated "metabolic" acidæmia resulting, *e.g.* from ingestion of NH_4Cl, ketosis of diabetes mellitus or starvation, retention of acids in nephritis.

B. The "metabolic" acidæmia has been compensated for; there is a fall in alveolar and arterial CO_2 tension and in H_2CO_3 concentration (cf. Fig. 58, B, C and D).

C. Uncompensated "metabolic' 'alkalæmia resulting, *e.g.* from ingestion of $NaHCO_3$ or potassium citrate, or vomiting.

D. The "metabolic" alkalæmia has been compensated for; there is a rise in alveolar and arterial CO_2 tension and H_2CO_3 concentration.

II. In "*respiratory*" changes of blood reaction the *change in H_2CO_3 is primary*.

E. Uncompensated "respiratory" acidæmia resulting, *e.g.* from breathing air rich in CO_2, increased formation of CO_2 as in exercise or failure to eliminate CO_2 because of disorders of breathing or of the circulation.

F. The "respiratory" acidæmia has been compensated for; there is an increase in plasma $BHCO_3$.

G. Uncompensated "respiratory" alkalæmia resulting from overventilation at rest.

H. The "respiratory" alkalæmia is compensated for; there is a decrease in plasma $BHCO_3$.

The legend to Fig. 57 should be carefully studied. The subject is considered further below.

Alkalæmia.—In the following conditions there is a tendency to alkalæmia (Fig. 57, G, H; C, D).

(1) *Excessive breathing at rest, i.e.* breathing in excess of the amount necessary to eliminate the CO_2 formed by metabolism (*respiratory alkalæmia or alkalosis*); see Figs. 56, 57, G, and p. 406.

(2) *Increase in the bicarbonate content* of the blood disturbing the ratio $BHCO_3/H_2CO_3$ to the alkaline side (Fig. 57, C) (*metabolic alkalæmia* (or *alkalosis*)). This type of alkalæmia occurs in the following conditions:

(i) Ingestion of bicarbonate or other alkali-producing salts, *e.g.* potassium citrate (p. 98).

(ii) A group of conditions associated with *vomiting*[1]: (*a*) pyloric obstruction (p. 105); (*b*) high intestinal obstruction (p. 108); (*c*) in some cases of infantile diarrhœa and vomiting. As explained on p. 105 there is usually a greater loss of Cl' than of Na^+ in the vomit; as a result plasma Na^+ previously bound with Cl' is freed and combines with HCO_3' to form $NaHCO_3$; alkalæmia therefore results.

It should be noted, however, that an increase in plasma $BHCO_3$ may occur in states of *acidæmia* which result from CO_2 retention. Thus in exercise, or when a CO_2-rich mixture is breathed, the CO_2 tension in the blood rises; secondly, by reactions with hæmoglobin, $BHCO_3$ also rises. The rise of H_2CO_3, which is *primary*, is relatively greater than the rise of $BHCO_3$, which is secondary; the ratio $BHCO_3/H_2CO_3$ is disturbed to the *acid* side. The condition is one of *respiratory acidæmia* (or *acidosis*) and as pointed out is associated with an increase of $BHCO_3$ (Fig. 57, E).

Symptoms of Alkalæmia.[2]—These are well seen clinically when excessive doses of alkali are given in the treatment of peptic ulcer. There is loss of appetite, headache, irritability and other changes in the mental state, nausea and vomiting; there may be aching pains and twitching in the muscles, flushing or perspiration and weakness which may pass into a state of prostration. Later tetany develops (cf. p. 1007); the breathing is slow, the pulse is rapid, and fits may occur. The urine volume varies (it may be markedly reduced); it is usually alkaline in reaction and *renal damage* may be indicated by the presence of albumin, granular, and hyaline casts, and red and white blood cells.[3] [Sometimes, however, the urine may be *acid*, conserving base to maintain the crystalloid osmotic pressure while "sacrificing" the "needs" of blood reaction (p. 107).] The plasma bicarbonate is elevated, *e.g.* up to 90 c.c.%. The renal insufficiency which is present leads to an increase in the concentration of urea and other non-protein nitrogenous constituents of the plasma. Death may result in coma, apparently mainly as a result of uræmia. Post-mortem, the renal tubules may show marked degenerative changes.

Treatment consists in discontinuing the administration of alkali and giving NaCl and possibly also acidifying salts like NH_4Cl (p. 98) and NaH_2PO_4.

Acidæmia.—Acidæmia may be present owing to a diminution of the bicarbonate content or to an increase in the H_2CO_3 content of the blood (Fig. 57, A, B; E, F).

(1) *Primary Reduction of $BHCO_3$.*—In the following conditions the *bicarbonate content of the blood is lowered*, as it is used to buffer acids which have accumulated in the body (*metabolic acidæmia* (*acidosis*)) (Fig. 57, A). The usual responses to acidæmia develop: increased pulmonary ventilation,

[1] Burnett *et al.*, *J. clin. Investig.*, 1950, *29*, 169, 175.
[2] Cope, *Clin. Sci.*, 1936, *2*, 287.
[3] Berger and Binger, *J. Amer. med. Assoc.*, 1935, *104*, 1383.

lowered alveolar CO_2 tension, acid urine, and increased ammonium content are present.

(i) The ingestion of wood spirit, *i.e.* methyl alcohol, which gives rise to formic acid in the body.

(ii) Ammonium chloride after absorption is de-aminized in the body: NH_3 is split off and HCl is left, which enters the blood (p. 98).

(iii) *Diabetes Mellitus.*—There is an abnormal accumulation of β-hydroxybutyric and acetoacetic acids; they are neutralized by the plasma $BHCO_3$;

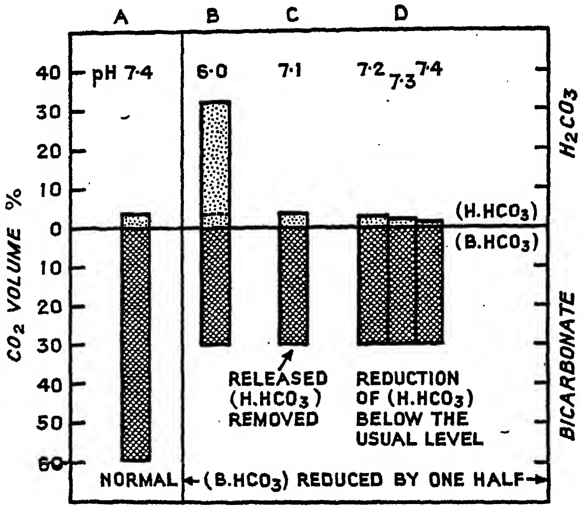


FIG. 58.—Relationship of Carbonic Acid–Bicarbonate Ratio in Plasma to *p*H of Plasma in Uncompensated and Compensated Acidæmia due to reduction of $BHCO_3$. (After Gamble, *Extracellular Fluid*, Harvard University Press, 1949.)

H_2CO_3 (stippled column) above horizontal line; $BHCO_3$ (hatched column) below horizontal line, in volumes of CO_2 per cent.
A. Normal $H_2CO_3/BHCO_3$; *p*H 7·4.
B, C, D. Reduction of plasma $BHCO_3$ to half normal value by addition of acid.
B. All the CO_2 which is released by the decomposition of $BHCO_3$ is retained in the plasma as H_2CO_3 resulting *p*H is 6·0.
C. All the CO_2 released by the decomposition of $BHCO_3$ is "blown off" by increased pulmonary ventilation so that H_2CO_3 remains normal (but $BHCO_3$ is halved); resulting *p*H is 7·1.
D. All the released CO_2 is blown off and in addition, by a further increase in the pulmonary ventilation the H_2CO_3 is progressively reduced below its normal value; *p*H progressively returns to normal.

the ratio $BHCO_3/H_2CO_3$ is disturbed to the acid side because $BHCO_3$ is reduced and H_2CO_3 is increased. Fig. 58 illustrates the sequence of events when half the $BHCO_3$ is used to neutralize the acids. In the absence of the vital reactions the *p*H would alter from 7·4 to 6·0. Breathing is, however, stimulated; it first eliminates all the excess CO_2 formed by the buffering reaction in the blood; it then turns out still more CO_2, thus reducing the alveolar CO_2 tension, and therefore the arterial CO_2 tension to *below* normal the ratio of $BHCO_3/H_2CO_3$ is restored approximately to normal. The kidney then comes into action; a highly acid urine is passed; large amounts of NH_3 are formed to neutralize the acids; fixed base (*i.e.* Na^+) is reabsorbed helping to restore and preserve the plasma $BHCO_3$ (see also p. 924).

(iv) Similar changes on a smaller scale occur in starvation (p. 901).

(v) In nephritis (p. 76) and uræmia (pp. 77, 456).

(vi) In cases of infantile diarrhœa and vomiting in which the loss of base (cations) in the fæces exceeds that of acid radicals (anions).

(2) *Primary excess of CO_2 (in the form of H_2CO_3) in the blood (respiratory acidæmia* or *acidosis*) occurs (Fig. 57, E, F)—

(i) In experimental breathing of CO_2-rich mixtures (p. 394).

(ii) *In Muscular Exercise.*—The acidæmia of muscular exercise is discussed on pp. 436 *et seq.* In moderate exercise it is due solely to excess CO_2 formed by the active tissues. In severe exercise lactic acid enters the blood stream, decreasing the plasma bicarbonate and so producing a " metabolic " acidæmia as well.

(iii) During *sleep* the alveolar CO_2 rises because of inadequate pulmonary ventilation (p. 405); the urine is highly acid.

(iv) In poisoning with morphine and other narcotics and anæsthetics, owing to depression of the respiratory centre. The pulmonary ventilation becomes inadequate and so CO_2 is not eliminated in sufficient amounts.

(v) In heart failure when the pulmonary epithelium may be damaged so that CO_2 cannot diffuse out readily (p. 461).

(vi) In emphysema, where the CO_2 tension in the alveolar air is raised. This condition is partially but not fully compensated for by an increased bicarbonate content of the blood (cf. p. 459).

(vii) In obstruction to the main respiratory passages (p. 458).

In the whole of this group there is, as explained (p. 99), a secondary increase of $BHCO_3$; the acidæmia persists as long as the $BHCO_3$ does not increase to the same extent as the H_2CO_3.

LABORATORY INVESTIGATION OF ACIDÆMIA AND ALKALÆMIA.—In these conditions the determination of the plasma bicarbonate is often useful (pp. 90, 92). If it is decreased there is either " metabolic " acidæmia or " respiratory " alkalæmia; if it is increased there is either " metabolic " alkalæmia or " respiratory " acidæmia. Usually the right conclusion can be drawn from a consideration of all the clinical data. If doubt persists then the pH of the blood must be directly determined.

EFFECTS ON BODY FLUIDS OF OBSTRUCTION OF THE ALIMENTARY CANAL [1]

General Considerations.—To survive, the body needs oxygen, water, electrolytes, and food in appropriate amounts. In intestinal obstruction there is loss of water and electrolytes in varying proportions, thus disturbing both the volume and composition of the body fluids; the fluid which is lost from the body may be neutral, acid, or alkaline in reaction so leading to alkalæmia or acidæmia. The failure to retain food is of little importance from the short-term point of view but becomes increasingly significant as the period of obstruction is prolonged. If the obstruction is accompanied by changes in the state of the intestinal blood vessels there may be loss of

[1] Gamble, *Extracellular Fluid,* Harvard University Press, 1949. Cooper, *Arch. Surg.,* 1928, *17,* 918. M'Iver, *Intestinal Obstruction,* N.Y., 1934. Holt, *Lancet,* 1939, ii, 61.

plasma or even of whole blood from these vessels into the lumen of the gut increasing the difficulties of plasma volume regulation. The clinical condition is aggravated by pain, vomiting, and the development of infection.

The results of simple salt deprivation and of water deprivation (which produces water and salt loss) are described on pp. 64, 66. As mentioned on

Secretion.	Volume c.c.	Total NaCl g.
Saliva . . .	1500	7
Gastric Juice . .	2000–3000	15–30
Bile . . .	500–1000	7
Pancreatic Juice . .	500–800	8
Succus Entericus . .	? 3000	? 18

p. 6, an immense water and electrolyte turnover takes place in the alimentary canal. The Table above indicates the daily average volumes of the digestive juices which enter the bowel from the blood and which are normally practically completely reabsorbed into the blood.

Fig. 59 sets out the electrolyte composition of the main gastro-intestinal secretions and compares it with that of plasma.

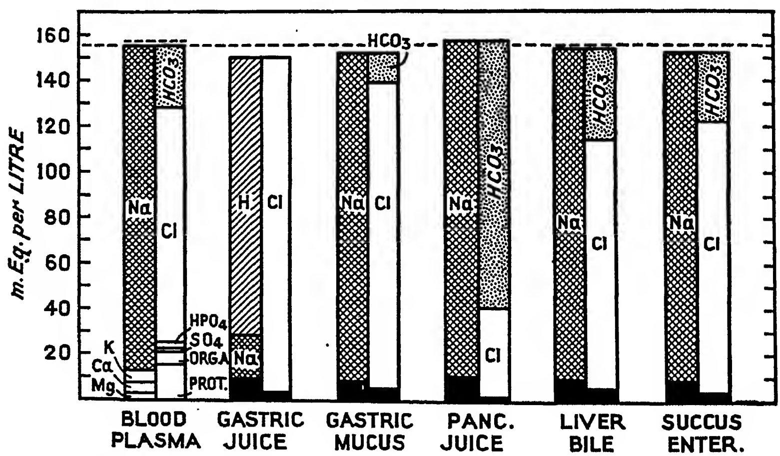


FIG. 59.—Electrolyte Composition of Plasma, Gastric Juice, Gastric Mucus, Pancreatic Juice, Liver Bile and Succus Entericus. (After Gamble, *Extracellular Fluid*, Harvard University Press, 1949.)

Complete loss of the gastro-intestinal secretions for 24 hours would deprive the body of about 8 L of water or nearly one-fifth of the total body water, a loss sufficient in itself to produce a grave clinical condition. This water loss would be accompanied by the loss of the electrolytes in the juices, the net result being the loss of an approximately isosmotic solution of varying acid-base balance. The total salt loss in a day might be 50 g. compared with a total salt content in all the body fluids of the order of 100 g. (and a

daily salt intake of say 10 g.). It should be carefully borne in mind that over and above the obvious loss by vomiting, unavoidable fluid losses are taking place: in the lungs (400 c.c.), by insensible perspiration (600 c.c.), and in the urine (the minimum volume necessary to eliminate the waste products from the body is about 500 c.c.), or a total of at least 1500 c.c. If sweating is taking place the water loss is even greater. If none of the water drunk is retained then the water lost is the volume vomited *plus* the loss by the other routes just indicated (minus the water resulting from oxidation of the food-stuffs). The loss of body fluid in intestinal obstruction is not a question of simple dehydration, *i.e.* simple loss of water; it is a loss of water *accompanied by loss of electrolytes.*

Results of Loss of Gastric Secretion.—The main gastric glands (p. 775) secrete a fluid which, apart from its enzyme content, is practically a 0·1 N. solution of HCl; gastric mucus and the secretion of the pyloric region is alkaline, the contained electrolytes being Na^+, Cl', and HCO_3'. The volume of fluid secreted by the main gastric glands normally far exceeds that formed in the pyloric glands. Loss of gastric secretion (by vomiting or aspiration) leads to loss of water and of Cl', a relatively smaller loss of Na^+ and a variable loss of H^+ (acid). In complete pyloric obstruction none of the water drunk can be retained. It must be remembered too that the unavoidable water loss is going on all the time.

(1) Experimental Pyloric Obstruction.—In a rabbit with pyloric obstruction the loss of water and electrolytes in 36 hours was as follows:

Water lost, 203 c.c. [initial plasma volume, 83 c.c.].
Changes in Na^+ and Cl' (in m.Eq.) were:
Na^+ lost, 27 m.Eq. [initial *total* plasma Na^+: 14 m.Eq.].
Cl' lost, 30·9 m.Eq. [initial *total* plasma Cl': 8·5 m.Eq.].

As the loss of water, Na^+, and Cl', greatly exceeded that present initially in the whole plasma it is clear that the interstitial fluid must have been drawn upon to a considerable extent. As the Cl' lost is greater than the loss of Na^+, some of the extracellular Na^+ is freed to combine with free CO_2 (as HCO_3') to form $NaHCO_3$, thus producing an *alkalæmia.*

(2) Clinical Pyloric Obstruction.—In clinical pyloric stenosis the same fundamental changes occur. If the composition of the total fluid lost (by all routes) were isosmotic, then the crystalloid o.p. of the extracellular fluid would remain unchanged (though its volume is diminished). Under these circumstances there would be no withdrawal of intracellular fluid. In practice the net result clinically is a *relatively greater loss of electrolytes than of water.* The extracellular fluid thus becomes *hypotonic* as well as reduced in volume; the plasma shows a marked fall of Cl', a less marked fall of Na^+, and an increase of HCO_3' (as $NaHCO_3$) (Fig. 60). Because the crystalloid osmotic pressure of the extracellular fluid is reduced, fluid flows into the cells, thus further decreasing the extracellular fluid volume. The volume of the *plasma* is, however, partly protected by the rise of plasma *protein* osmotic pressure which helps to retain fluid in the blood (note the increase in R in Fig. 60) which is due mainly to increased protein concentration. As the plasma volume decreases, the red cell count and the hæmoglobin concentration correspondingly rise. The kidney responds as expected by a reduction in

the volume of urine to the minimal value, by the maximal degree of reabsorption of electrolytes and by the secretion of an alkaline urine. Ultimately the kidney fails because of impaired blood supply and the harmful effects of the electrolyte imbalance (chiefly lack of Na^+ ions) (cf. p. 64).

(3) Gastric Aspiration Plus Water Replacement.—Fig. 61 illustrates an experiment performed in man. Continuous aspiration of the gastric juice was carried out for several days, *water* and *energy* requirements being made good by drinking water and a glucose solution. The loss is thus of electrolytes only: of Cl′, relatively less Na^+, and of H^+ (acid), *i.e.* there was a depletion

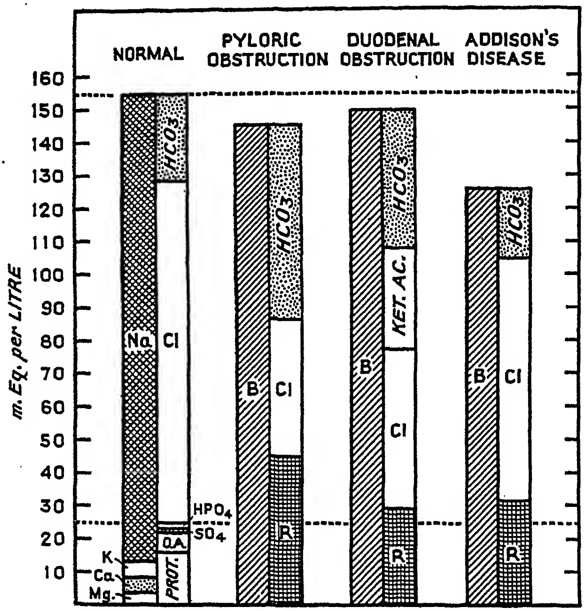


Fig. 60.—Electrolyte Composition of Normal Plasma and the Plasma in Pyloric Obstruction, Duodenal Obstruction, and Addison's Disease. (After Gamble, *Extracellular Fluid*, Harvard University Press, 1949.)

O.A.=organic acids; B=total base; Ket. ac.=Keto-acids.
R=combined concentration of organic acid, phosphate, sulphate, and protein.
Note increase of R in all three pathological conditions. In "Duodenal Obstruction" the keto-acids have been depicted separately; to get the comparable value of R the keto-acid concentration should be added to the R shown.

of only salt and acid. The extracellular fluid became hypotonic (though unchanged in volume) and water doubtless flowed into the cells, increasing their volume. The plasma showed the expected decrease in Cl′ and an increase in $BHCO_3$ (the plasma bicarbonate rose from 65 to 140 c.c. of CO_2 per 100 c.c.); there was no hæmoconcentration, showing that the water intake was adequate. The kidney showed the characteristic failure which results from Na^+ and Cl′ depletion: the volume of urine was reduced to only 500 c.c. daily in spite of a plentiful water intake; although the urine was alkaline (as would be expected in alkalæmia) the plasma bicarbonate continued to rise. Nitrogenous excretion was inadequate (*e.g.* the total urea output was

decreased) leading to nitrogenous retention (azotæmia): the blood urea rose from 30 mg. to 130 mg. per 100 c.c. On administration of saline the plasma Cl′ rose to normal, replacing HCO_3'; the plasma bicarbonate fell and plasma Na^+ doubtless rose. Renal activity was restored and the retained nitrogenous constituents were soon eliminated, blood urea returning to normal.

Similarly, if a *salt-free* solution (*e.g.* a glucose solution) is injected in clinical cases of pyloric obstruction in sufficient volume to make good the water loss, the plasma volume can be fully restored and the anhydræmia relieved, but the alkalæmia and (to a smaller extent) the renal failure are unaffected.

Acid Urine in Alkalæmia from Salt Loss due to Vomiting.—In some patients an acid urine is passed in spite of the alkalæmia. The explanation may be that plasma crystalloid osmotic pressure is being sustained at the " expense " of blood reaction regulation. The excretion of an *alkaline* urine involves the loss of base as bicarbonate ($BHCO_3$) (Fig. 52); in addition extra base is lost because the urinary phosphate is excreted mainly in the form of B_2HPO_4 instead of BH_2PO_4. The loss of this extra base lowers the $BHCO_3$ content of the plasma, but its crystalloid osmotic pressure also falls; secretion of an alkaline urine in states of salt lack thus preserves blood reaction at the " expense " of crystalloid osmotic pressure. On the other hand, if an *acid* urine is excreted, there is no loss of base in the form of $BHCO_3$ or of B_2HPO_4; the sulphate and even the phosphate may be excreted in the urine as ammonium salts. As a result the plasma $BHCO_3$ is maintained preserving (as stated above) crystalline osmotic pressure at the " expense " of blood reaction.

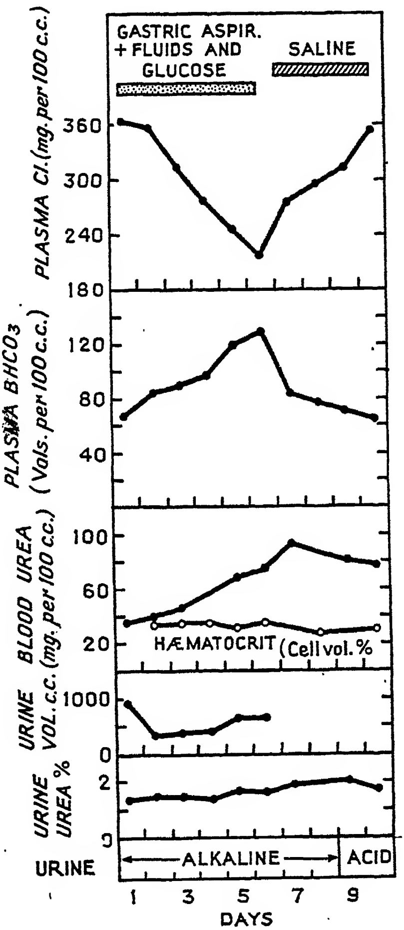


FIG. 61.—Changes in Blood and Urine following Experimental Aspiration of Gastric Contents. (Nicol, *Quart. J. Med.*, 1940, *9*, 98.)

Adult in clinically normal state (apart from slight alkalæmia due to alkali therapy). Aspirate 1½–2 litres of gastric contents daily for 5 days; salt intake restricted; fluid and glucose given freely. Note fall in plasma Cl′, rise in blood urea and plasma bicarbonate and decreased urine volume. On 6th day give saline injections: blood composition rapidly restored to normal.

If NaCl is given to such cases, a highly alkaline urine containing large amounts of $BHCO_3$ and B_2HPO_4 is passed and the blood reaction rapidly

returns to normal. The explanation is as follows: in salt loss due to vomiting the fall of plasma Cl′ may be greater than that of Na^+; the plasma Na^+ is thus restored to normal before the Cl′. As soon as there is a slight excess of Na^+ it is excreted as $NaHCO_3$, thus restoring the blood reaction.

Alkalæmia from Vomiting in Cases of Gastric Achlorhydria.—Cases of carcinoma of the stomach in which *no gastric HCl* is being secreted may develop alkalæmia if severe vomiting takes place. As Na^+ and Cl′ are being lost in equivalent amounts the question arises where the Na^+ comes from with which the extra plasma HCO_3' is combined. There is no certain answer, but possibly in such cases disturbed renal function may lead to more efficient reabsorption of Na^+ than of Cl′; the surplus Na^+ unites with HCO_3' to form bicarbonate, so producing an alkalæmia.

Simple High Small-Intestine Obstruction.—With simple obstruction high up in the small intestine, absorption of saliva, gastric juice, bile, and pancreatic juice (and ingested food and water) is interfered with; these fluids accumulate above the block and distend the bowel, setting up powerful, colicky peristaltic movements and profuse vomiting. As is evident from the data in the Table on p. 104, in a very short time an enormous loss of water may occur which is probably the main factor responsible for death. It should be borne in mind that the usual additional fluid is being lost from the body in the urine, from the skin (insensible perspiration 600 c.c. and sweat (p. 467)), in the breath (400 c.c.), and in the fæces (if any are passed).

The fluid lost is approximately isosmotic and neutral; as water is being lost at the same time by the uncontrollable routes there may be relatively more loss of water than of salt; on the other hand, if some ingested water is being retained the loss of salt may be the greater. Examination of the blood (Fig. 60) suggests that *salt loss exceeds that of water*, the changes observed being essentially the same as in pyloric obstruction, *i.e.* (i) fall of plasma Cl′ (*hypochlorœmia*); (ii) *alkalœmia* from increased plasma bicarbonate due to Cl′ loss exceeding that of Na^+, the plasma Na^+ being thus freed to combine with free CO_2 (HCO_3') to form bicarbonate; (iii) nitrogenous retention (*azotœmia*) secondary to renal failure. The urinary Cl′ disappears owing to complete absorption of Cl′ in the tubules; the urinary reaction is usually alkaline but may be acid (p. 107).

The intestinal distension referred to above aggravates the condition in various ways: (i) intestinal secretion is stimulated; (ii) vomiting is worsened; (iii) there is interference with the circulation in the gut wall leading to venous and capillary engorgement. The rise of capillary pressure and the altered permeability of the capillary endothelium leads to the escape of large quantities of protein-rich fluid into the gut wall. When half the length of the small intestine is involved a volume of fluid equal to half the plasma volume may be lost in this way. All the factors enumerated above intensify the anhydræmia. Signs of dehydration make their appearance when the net fluid loss is equal to 6% of the body weight, *i.e.* to about 4000 c.c.; from the calculations given on p. 104, it is obvious that such a loss can occur with startling rapidity.[1]

The swollen, overstretched intestinal wall loses tone and contractility, and finally complete paralysis develops above the level of the obstruction; hæmorrhages may occur in the gut wall, patches of necrosis develop, and

[1] Aird, *Brit. J. Surg.*, 1938–39, *26*, 418.

finally perforation may take place. At any stage infection may complicate the issue. In advanced cases respiration is depressed owing to the alkalæmia and periods of apnœa may occur; there may be signs of *tetany*. The plasma volume is as usual initially well maintained because the fluid loss is in large measure borne by the interstitial spaces; but as the anhydræmia becomes worse the plasma volume is substantially decreased, and circulatory failure occurs, producing anoxia and cyanosis and aggravating the renal failure.

LOW SMALL-INTESTINE OBSTRUCTION.—Symptoms develop more gradually and are less severe as absorption of water and solutes can take place more satisfactorily above the level of the block. On the other hand, as intestinal distension becomes more marked, local circulatory disturbances develop; loss of fluid into the gut wall itself then becomes a serious factor, intensifying the anhydræmia.

Principles of Treatment of Water and Electrolyte Lack.—The essential principle is to determine accurately the nature of the disturbance and to take the necessary measures promptly to correct it. In clinical conditions in which varying degrees of depletion of water, Na^+ and Cl' and *p*H changes have produced a complex alteration of the electrolyte pattern and the fluid volume and distribution, it is found that *as soon as enough water and salt (sodium chloride) have been provided to restore renal activity the kidneys can be relied upon to restore the electrolyte pattern of the body fluids accurately to normal by their own specialized activities.*

(1) SIMPLE WATER LOSS.—When practicable give plenty of water to drink; in addition inject 5% glucose solution intravenously to produce rapid effects: the sugar is metabolized (yielding energy) leaving the water of the solution to make good the water deficiency. Salt (or weak saline) should also be given to make good the associated salt lack (p. 67).

(2) WATER AND SALT LOSS.—Saline (0·9%) is given by all convenient routes until *Cl′ appears in the urine,* indicating that the Cl′ content of the extracellular fluids is restored to normal. Water or a glucose solution can then be given to make good any remaining shortage of body water.

(3) SIMPLE INTESTINAL OBSTRUCTION.—It is clear from what was stated on p. 106 that administration of water or glucose solution is useless; a bicarbonate solution aggravates the pre-existing alkalæmia. *Saline* (0·9%) must be given on the lines indicated in (2). Tetany can be relieved by acidifying procedures, *e.g.* inhalation of 5% CO_2 or injection of soluble calcium salts (*e.g.* 5% $CaCl_2$). Great relief is obtained clinically by continuous aspiration of the gastric and duodenal contents to relieve the vomiting and the intestinal distension. It should be remembered however that continuous aspiration on these lines, *e.g.* for 6 days, may withdraw over 90 g. of sodium chloride from the body and a very large volume of fluid. The water and salt loss so produced is just as real as when it results from vomiting, though less spectacular and less distressing.

A salt-lactate solution (containing NaCl and Na lactate) is of value in states of low plasma bicarbonate; the addition of KCl may be useful.[1] The lactate ion is oxidized freeing the Na^+ to combine with CO_2 to re-form $NaHCO_3$

[1] Butler's solution contains the following constituents in m.Eq./L: Na^+, 30; K^+, 15; Cl', 22; Lactate′, 20; also HPO_4'', 3 mg/L; all in 5% glucose (cf. p. 971).

(as occurs during the recovery phase after severe muscular exercise (p. 440).

Isotonic (1·5%) sodium bicarbonate solution is of value as an emergency measure in dangerous degrees of *acidosis* due to $NaHCO_3$ depletion until salt or salt-lactate solutions can be given.

Obstruction initially associated with Interference with Local Blood Supply (Strangulation).—Usually the veins are obstructed. Capillary engorgement becomes intense; the peritoneum is filled with an exudate resembling plasma; the gut lumen is filled with a thick red exudate. Organisms proliferate rapidly in the contents of the bowel and in its wall liberating *toxins* which are probably the cause of death. The clinical condition is aggravated by serious loss of plasma or of blood into the bowel and the presence of infection. After a time the fluid in the bowel becomes darker and finally black and exceedingly toxic: if injected intraperitoneally into another animal it proves rapidly fatal.

If the terminal part of the ileum is anastomosed to the rectum and the intervening isolated portion of the colon is closed, no evidence of intoxication appears. The lower end of the closed colon is later found filled with a sausage-like mass having the appearance and consistency of normal fæces. No toxins are produced and no symptoms develop.

ŒDEMA

Œdema.—By œdema is meant an abnormal accumulation of fluid in the intercellular spaces; it occurs *physiologically* in muscles during, and for some hours after, activity. For clinically demonstrable œdema to occur in the skin and subcutaneous tissues the volume of extracellular fluid must increase locally by about 10%.

Rôle of Abnormal Capillary Permeability.[1]—Increased capillary permeability to water and dissolved crystalloids *speeds* up interchanges between the plasma and the tissue spaces but does not alter the *final* distribution set up. If, however, the capillaries become excessively permeable to protein, plasma protein in varying amounts escapes from the plasma into the tissue spaces and the normal balance of forces is disturbed. As the protein concentration in the plasma falls the holding power of the blood for fluid is decreased and so more tends to escape; at the same time the protein content and osmotic pressure of the tissue-spaces rise leading to greater suction of fluid into the tissue spaces and its retention there in excessive amounts, causing œdema. The œdema or wheal due to skin injury and injected histamine is produced in this way (p. 324). In all forms of œdema the accumulation of fluid in the tissue spaces leads to a *rise of tissue fluid pressure* which tends to check further escape of fluid from the blood vessels.

In susceptible subjects, *light* stroking of the skin may give rise to the flare and wheal (*urticaria factitia*, Fig. 196); the injection of histamine into the skin in these cases produces a local response of normal character, indicating that there is no state of unusual reactivity of the local vessels. It would appear that the precursor of histamine, which is present in the skin cells, is *excessively unstable* in this disorder, so that very trivial injuries are sufficient to cause the liberation of histamine in sufficient concentration to give the full "triple

[1] Normal capillary permeability is discussed on p. 6; see also pp. 17–21.

response." Various poisons taken in the food, or elaborated in the body, may also give rise to urticaria, possibly by liberating histamine in the skin (cf. p. 337).

The pathological varieties of œdema will now be discussed.

Causes of Œdema.—1. **Venous Obstruction.**—Several factors are probably involved: (i) Venous obstruction raises venous and capillary pressure and thus increases filtration of fluid out into the tissues. (ii) In cases in which the venous obstruction is associated with inflammation of the vein wall (phlebitis) or of the surrounding tissue, reflex *arterial spasm* may occur leading, when the lower limb is involved, to a " white leg ": the skin is pale and cold and the pulse in the distal arteries is diminished or absent. (iii) As a result of the changes in the arteries and veins the blood-flow to the affected part is slowed down, leading to local accumulation of vasodilator metabolites which may increase capillary permeability and so permit the escape of a fluid excessively rich in protein. Some of the metabolites may be osmotically active and thus further increase tissue-fluid formation and retention. The extent to which these various factors operate determines the degree of œdema produced.

2. **Cardiac Œdema.**[1]—In congestive heart failure œdema may be a very prominent symptom. The mechanisms concerned are *vascular* and *renal*.

(1) The *venous pressure* in the median basilic vein in severe cases of congestive failure is raised from the normal level of 10–15 cm. H_2O to 30–35 cm. H_2O. The capillary pressure must be correspondingly higher, and fluid therefore tends to filter out into the tissue spaces. The depressed cardiac output fails to maintain an adequate blood supply to the organs; cellular metabolism is disturbed and metabolites instead of being rapidly removed accumulate in the tissue spaces; the local osmotic pressure rises and retention of interstitial fluid follows. The œdema fluid contains protein in concentrations varying from a trace to 0·5% indicating some, but on the whole not greatly increased permeability of the capillaries to protein. The œdema fluid is most readily formed in the dependent parts because there the capillaries are subjected to the additional hydrostatic pressure of the column of blood (p. 17). The œdema is noticed in the ankles in the evening if the patient is walking about, and passes off during the night.

(2) A *renal* factor is frequently in operation also. In a normal person the sodium excretion in the urine approximately equals the intake, thus keeping the Na^+ concentration in the plasma and interstitial fluid within normal limits. In congestive heart failure, urinary Na^+ excretion is reduced, even down to 1% of normal (Fig. 62). Owing to Na^+ retention, water is secondarily retained in the extracellular fluid to preserve the normal osmotic concentrations; the flow of urine is corresponding decreased. The reasons for the failure of Na^+ excretion are obscure, but the following suggestions have been made.

(i) Owing to a reduced renal blood flow, the glomerular filtration rate is decreased leading secondarily to retention of sodium. In some cases of congestive heart failure with marked œdema the measured renal plasma flow was 20–30% of normal, and the glomerular filtrate volume 30–50% of normal;

[1] Brod and Fejfar, *Quart. J. Med.*, 1950, *19*, 187, 221. Merrill, *J. clin. Investig.*, 1946, *5*, 389; *Bull. N.Y. Acad. Med.*, 1948, *24*, 607.

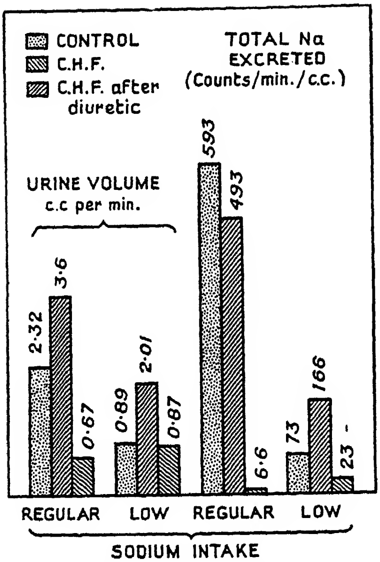


FIG. 62.—Sodium and Water Output in Congestive Heart Failure. (Reaser and Burch, *Proc. Soc. exp. Biol. Med.*, 1946, *63*, 543.)

The block diagrams compare the urine volume (in c.c. per minute) and sodium excretion in a normal subject and a patient with congestive heart failure before and after administering a mercurial diuretic.

C.H.F.: congestive heart failure.

Regular and Low: normal diet and salt-free diet respectively.

Before beginning the observations both the normal subject and the patient were given 200 c.c. of a 1% solution of $^{22}NaCl^*$ (*i.e.* sodium chloride containing radio-active $^{22}Na^*$, half-life=3 years) which diffuses throughout the body fluids; the rate of excretion of $^{22}Na^*$ in the urine is determined with a Geiger counter and is directly proportional to total Na excretion.

Urine volume: On a *normal* salt intake the normal subject secretes 2·32 c.c. of urine per minute compared with 0·67 c.c. in the patient; the mercurial diuretic greatly increases urine flow in the latter (to 3·6 c.c.).

On the *low* salt intake the urine volume is equally low in both subjects but is again increased in the patient by the diuretic (from 0·87 to 2·01 c.c.).

Na excretion: Note the failure of Na excretion in the patient and the enormous increase (from 6·6 to 493 counts on a normal diet, and from 23 to 166 on a salt free diet) produced by the diuretic.

the filtration fraction was 0·3–0·5% which is suggestive of constriction of the efferent glomerular vessels (p. 25).[1] But, as explained on p. 73, the glomerular filtrate must fall to negligibly low levels before failure to excrete the normal Na^+ intake can occur; such a degree of renal failure would be inevitably associated with marked retention of nitrogenous waste products, which in fact does not take place.

(ii) More probably, there is a primary derangement of *tubular* function expressing itself by an excessive degree of Na^+ reabsorption, *i.e.* more Na^+ than normally is passed back into the blood from the lumen of the tubules; the same mechanism may be responsible for excess reabsorption of water. The tubular abnormality may be secondary to the local circulatory disturbance or it may result from some endocrine disorder, *e.g.* excessive secretion of the antidiuretic hormone or of adrenal cortex steroids (p. 967). Whatever the nature of the tubular disturbance it is annulled by the administration of the *mercurial diuretics* (Fig. 62). A *low salt intake* relieves the condition by arresting the accumulation of sodium in the body.

3. Renal Œdema.—Œdema occurs (i) in *Type* 2 *nephritis* (*nephrosis* or *nephrotic syndrome*) (p. 75); (ii) in the acute stage of *Type* 1 *nephritis* (p. 75); (iii) in *chronic nephritis* (p. 76).

(1) ŒDEMA OF TYPE 2 NEPHRITIS (NEPHROSIS).[2]—As explained on p. 75, the characteristic renal lesion is hyalinization of the basement membrane of the glomerular capillaries and lipoid accumulation in the renal tubules. In this disease the œdema is severe and is often accompanied by effusions into

[1] Normal renal plasma flow per minute is 700 c.c.; suppose it is reduced to a quarter, *i.e.* to 170 c.c. Normal glomerular filtrate volume per minute is 120 c.c.; suppose it is reduced to 40%, *i.e.* to 50 c.c. Filtration fraction (glomerular filtrate/renal plasma flow)$=\frac{50}{170}=0{\cdot}3$ (the normal fraction is $\frac{120}{700}=0{\cdot}17$).

[2] Leiter, *Medicine*, 1931, *10*, 135.

the serous cavities. In accounting for the œdema the following points have to be considered.

(i) There is a little direct evidence of general renal insufficiency: the urea clearance test (p. 39), urea concentration test (p. 69), ammonia formation (p. 97), blood nitrogen and hydrogen-ion concentration are all normal. Hypertension, cardiac enlargement, and uræmic symptoms do not develop.

(ii) The œdema fluid has a negligible protein content (about 0·1% compared with 0·6–1% in acute Type 1 nephritis); this indicates that general capillary permeability to protein is not increased.

(iii) The volume of urine is greatly reduced and may only amount to a few hundred c.c. daily. If water is ingested, or saline given intravenously, the excess fluid is excreted in the urine fairly rapidly and completely. From time to time spontaneous marked diuresis takes place. These observations suggest that the *primary trouble is not inability of the kidney to excrete water or salt*; the volume of urine is diminished because the fluid is retained in the tissue spaces.

(iv) The outstanding urinary abnormality is the excretion of large amounts of *protein* up to 5–20 g. per litre of urine daily, and numerous casts—hyaline, granular, or epithelial. The urinary protein is almost wholly *albumin* (and only to a slight extent globulin), and there is no detectable difference between it and the albumin of the plasma from which it doubtless comes. One must suppose that the permeability of the glomerular capillaries is altered so that the smaller albumin molecules escape from the plasma while the larger globulin molecules are retained.

(v) The expected changes are found in the plasma. The total plasma protein concentration is reduced, *e.g.* to 3·5–5·3%; there is no evidence that this is the result of an increased water content of the plasma. The protein depletion affects chiefly the albumin which falls from the normal 4% to 2 or even 1%; the albumin/globulin ratio is reduced from 1·5–2 to 0·5–1.[1] The globulin of the plasma may sometimes rise above normal owing to compensatory regeneration. The plasma protein concentration runs roughly parallel to the urinary protein loss, but other factors influence it too, *e.g.* varying degrees of fresh protein formation (which depends partly on the diet). The plasma protein osmotic pressure is greatly reduced. 1 g-% of albumin exerts an o.p. of 6 mm. Hg, while 1 g-% of globulin exerts an o.p. of only 1·5 mm. Hg.[2] The differential loss of plasma protein is most serious from this point of view, especially as the osmotic pressure per 1 g. falls as the total concentration falls. The normal balance between the blood and the tissues is deranged, and excessive amounts of fluid (and secondarily of salt) flow into the tissue spaces and are retained there. The fall of serum-albumin concentration and the associated lowering of the plasma protein osmotic pressure is the chief factor producing the œdema (Fig. 63) and the effusions into the serious cavities; owing to fluid retention in the tissues the urinary output is decreased. In nephrosis the critical level for œdema is said to be a total plasma protein concentration of 5%, and a serum albumin concentration of 2%. Attention must principally be paid to the latter.

(vi) *Salt deprivation* tends to decrease the œdema because water cannot be retained in the body unless an adequate amount of sodium chloride is

[1] For normal plasma protein values see p. 133. [2] Cf. p. 15.

available to give the normal crystalloid concentration. This observation is made use of in treatment.

(vii) There may be other subsidiary factors at work. There is a marked rise in the plasma *cholesterol* from 0·15–0·2 up to 1%, or even higher. The *basal metabolic rate* may be reduced to 20% below normal. These observations suggest that some tissue abnormality may possibly be present too.

Œdema and ascites of an identical kind can be produced experimentally in dogs by lowering the plasma protein concentration below 3% by the method of *plasmapheresis* (p. 137). This consists in removing whole blood

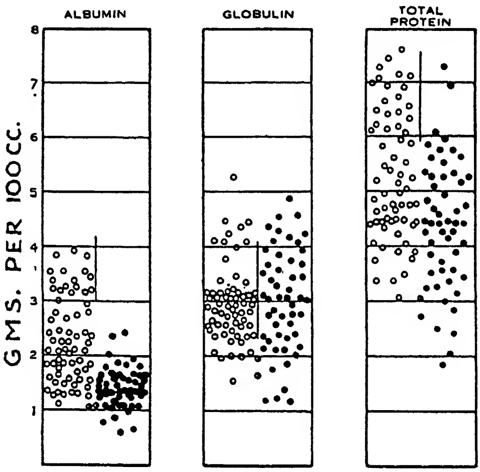


FIG. 63.—Relationship of Œdema to Lowered Plasma Protein Concentration due to Protein Starvation in Dogs. (Weech, *Bull. N.Y. Acad. Med.*, 1939, *15*, 74.)

Ordinate: concentrations of various proteins in plasma in g-%. Open circles indicate no œdema present; black circles indicate œdema present. Vertical lines in the middle of each column represent the range of normal variation. œdema is closely related to serum albumin concentration (œdema is commonly present when the albumin concentration is less than 2%), to a less extent to total plasma protein (when it is under 6%), and not at all to serum globulin concentration.

and reinjecting the corpuscles suspended in protein-free Ringer-Locke's solution. If this treatment is stopped for a day or two the protein concentration rises above 3% as a result of plasma protein regeneration and the œdema disappears.

A generous protein intake is advised for nephrotic œdema, the minimum being 1 g. protein per kg. body weight *over and above* the equivalent of the urinary protein loss. The ingested protein is of course not directly converted into plasma protein but provides the appropriate amino-acids required for large-scale protein regeneration. It must be remembered, too, that quite a small rise of plasma protein level may make all the difference between incapacity and comfort, by markedly modifying the degree of œdema. The full caloric needs of the patient should also be provided. Even the occurrence of some nitrogenous retention in the blood is not an indication for limiting the protein intake. Protein may perhaps also help by stimulating metabolism

and by giving rise to urea which acts as a powerful diuretic (p. 69); in fact, urea may be deliberately administered as a diuretic.

(2) Acute Type 1 Nephritis (Acute Glomerulo-nephritis).—The œdema occurs early in the disease and is slight and transient. It is suggested that the agent which is responsible for the inflammatory changes in the renal glomeruli also affects the capillaries generally. As in all forms of inflammation, capillary permeability is increased, leading to escape of protein into the interstitial spaces with consequent retention of fluid there. The protein content of the œdema fluid is 0·6–1%.

(3) Chronic Nephritis.—The œdema may be due to a number of factors, *e.g.* retention of electrolytes (Na^+, Cl') and water owing to deranged renal function and possibly increased capillary permeability.[1]

4. **Nutritional Œdema.**—During times of war and famine, when the amount of food available is insufficient, œdema (without albuminuria) develops in many members of the community. This occurred commonly in concentration and internment camps during the last war. The diet generally showed multiple deficiencies, being deficient in total calories (*e.g.* 1200–2000 Cal. daily), in total protein content (*e.g.* 20 g. daily) and in several of the vitamins (cf. pp. 1045 *et seq.*).

(1) Experimental Protein Deficiency Œdema.—Experimentally it is found that simple, prolonged, severe, protein deficiency produces loss of appetite, great weakness, and œdema; the last symptom is related to a great reduction of serum albumin and total plasma protein, and not at all to the concentration of serum globulin (Fig. 63). The œdema is attributed to the lowered plasma protein osmotic pressure (as in nephrosis, p. 113). If adequate amounts of protein are added to the diet, recovery ultimately occurs but improvement is slow owing to the poor appetite and perhaps also because of impaired liver function which leads to a slow rate of plasma protein synthesis (p. 137). As the plasma protein osmotic pressure rises the œdema disappears, and strength returns.

(2) Clinical Nutritional Œdema.—In the light of these results clinical nutritional œdema has been generally attributed to the deficient protein intake and the lowered plasma protein osmotic pressure; careful studies however show that several factors are involved.

(i) Blood examination showed the following average plasma protein concentrations in g–% in normal and malnourished groups of subjects:

Normal British		7·0 g–%
Internees:	without œdema	6·2 ,,
	with slight œdema	5·1 ,,
	,, moderate œdema	5·1 ,,
	,, severe œdema	4·8 ,,

The serum albumin in cases with œdema was 2·6–3·8%. In *individual* cases, however, the plasma protein or serum albumin level was *not* precisely related to the presence or the degree of œdema; in other words no critical plasma osmotic pressure level was found to be associated with œdema.

(ii) *Rest in bed* without any change in the diet has a remarkable beneficial effect. The œdema may disappear in 2–4 days; during this period the body weight may decline by 3–7 kg. owing to the elimination of excess

[1] The causes of *ascites* in *portal obstruction* are considered on p. 822.

extracellular fluid. The volume of urine passed in the first 48 hours may amount to ten litres; large amounts of Cl′ are excreted at the same time even up to 45 g. During this period, even on a full diet, the plasma protein may show no regular change. On getting up and doing some work, Cl′ and water are retained and the œdema returns.

(iii) The degree of clinical (*i.e.* subcutaneous) œdema is not always a reliable index of the amount of fluid retained because the fluid may be differentially stored in the internal organs, *e.g.* in the brain, mediastinal, perirenal or retroperitoneal tissues.

(iv) The astonishing effects of rest and exercise suggest that a *renal* factor is involved. It is possible that (as in congestive failure) there is some disorder in tubular function which is aggravated by exercise; one recalls that exercise calls forth a secretion of antidiuretic hormone (p. 53) which actively stimulates the reabsorption of water.

(v) The œdema which develops in the course of *marasmus* and *gastro-intestinal disorders of infancy* and in *cachectic* states may be due to similar mechanisms.[1]

Beri-beri Œdema.—In clinical *wet beri-beri* resulting from vitamin-B_1 deficiency, severe generalized œdema may be present; recovery occurs in 10 days on adding 2 mg. of -B_1 to the daily diet. This recovery also is associated with a greatly increased volume of urine. The factors responsible for the œdema of beri-beri have not been adequately analysed.

5. **Inflammatory Œdema** (see p. 222).

6. **Lymphatic Obstruction,** *e.g.* from growth, parasitic worms, may give rise to marked œdema which must be attributed to interference with the removal of tissue fluid. This is not altogether a satisfactory explanation, as under normal resting conditions the tissue fluid can be satisfactorily removed by the venous blood alone.

Pulmonary Œdema.[2]—Certain peculiarities of the pulmonary circulation must be borne in mind. In man, the normal pulmonary arterial diastolic pressure is on an average, 8 mm. Hg; the pressure in the pulmonary capillaries is 2–8 mm. Hg. Thus throughout the length of the pulmonary capillaries the blood pressure is substantially *less* than the osmotic pressure of the plasma proteins; there is consequently no tendency for fluid to escape from the pulmonary capillaries into the neighbouring tissues, including the air-sacs. As a result the air passages are normally kept dry except for the water vapour which is given off by the blood and saturates the respired air (p. 364).

Fluid can escape from the pulmonary capillaries into the air-sacs in two sets of circumstances:

(i) When the pulmonary capillary endothelium is damaged its permeability increases; plasma protein leaks out and attracts and holds a certain volume of excess fluid in the extravascular spaces including the air-sacs.

(ii) When the pulmonary capillary blood pressure exceeds the plasma protein osmotic pressure of 25 mm. Hg excess fluid is filtered out from the blood into the air-sacs. Sometimes both factors (i) and (ii) are simultaneously involved.

The fluid which enters the lungs is beaten up into a froth by the pulmonary ventilation and mechanically hinders the flow of air into and out of the depths

[1] See Keys *et al.*, *Biology of Human Starvation*, Minneapolis, 1950, vol. II, 921.
[2] Cameron, *Brit. med. J.*, i., 965. Cf. also p. 307.

of the lungs. If there is swelling of the pulmonary epithelium gaseous interchange, especially of oxygen, is interfered with; anoxia develops leading ultimately to respiratory arrest. The volume of fluid lost from the plasma into the lungs may in extreme cases be so large as to produce hæmoconcentration. Some of the œdema fluid drains away in the pulmonary lymphatics; but as the rate of flow along these vessels is slow they cannot cope with rapidly formed exudates. When the permeability of the pulmonary capillaries is increased, intravenous infusion of plasma (to make good the fluid loss) merely leads to a further rapid escape of fluid from the blood, aggravating the pulmonary œdema.

If *saline* is introduced into the lungs it is rapidly taken up into the blood owing to the suction action of the osmotic pressure of the plasma proteins. On the other hand, *plasma* introduced into the lungs is absorbed slowly, persisting for days. Similarly, *protein-rich* œdema fluid or inflammatory exudate in the air-sacs (such as is formed in pneumonia) clears up only gradually.

The common clinical and experimental causes of pulmonary œdema are as follows:

(1) Conditions associated with Increased Pulmonary Capillary Permeability.—(i) Inhalation of irritating gases like phosgene, mustard gas, or lewisite vapour, ammonia vapour, and many irritating smokes: all these act directly on the lung tissue with which they come in contact.

(ii) Absorption into the blood of chemicals like iodides, thiourea, or methyl salicylate which have a *selective* injurious action on the pulmonary capillaries.[1]

(iii) *Neurogenic Œdema.*—Pulmonary œdema commonly occurs as a terminal event in lesions of the central nervous system. Experimentally raised intracranial pressure may induce an œdema which is partially or wholly annulled by section of the vagi. It is suggested that changes in the central nervous system may lead to a discharge of impulses along nerve paths (possibly the vagi) to the pulmonary capillaries increasing their permeability and so leading to the escape of a protein-rich œdema fluid.

(2) Circulatory Failure.—(i) In *chronic congestive heart failure* (p. 460) the left auricular pressure rises and as a result the pressure in the pulmonary veins and the capillaries also rises, driving more fluid out into the air-sacs. It should be remembered that these pulmonary vessels are very distensible so that a large increase in their blood content must occur before much change of pressure takes place. As the œdema fluid is rich in protein there is probably some associated increased capillary permeability. Experimentally it is found that when there is some pulmonary capillary damage quite a moderate rise of capillary pressure may produce marked œdema.

(ii) Acute œdema of the lungs occurs in *acute left ventricular failure* associated with temporary normal activity of the right ventricle leading to pulmonary congestion and raised capillary pressure (p. 464). In some patients with circulatory disorders, the onset of a severe bout of *coughing* may likewise induce pulmonary œdema by obstructing the venous outflow from the lungs (p. 463). In grave cases both lungs are entirely filled with

[1] Certain agents act selectively on capillaries elsewhere; *e.g.* para-phenylene-diamine produces œdema of the head and face; œstrogens act on the sexual skin in certain monkeys.

œdema fluid, and frothy, possibly blood-stained, fluid may pour out of the mouth.

(3) Massive Intravenous Infusions of Saline may induce pulmonary œdema. If there is an associated rise of pulmonary venous pressure the œdema is presumably due to this cause; but the œdema may occur in the absence of this factor. The explanation is then unknown.

THE CEREBROSPINAL FLUID [1]

The cerebrospinal fluid is considered here because it is a subdivision of the extracellular fluid. As examination of the cerebrospinal fluid is an important routine of clinical medicine, it is essential to have a clear conception of its physiology and anatomy.

Anatomy.—The central nervous system is enveloped by the meninges; from without inwards these are termed dura, arachnoid, and pia. The dura consists of fibrous tissue lined by endothelium. Separating it from the arachnoid is the subdural space, which contains a small amount of fluid resembling lymph (Fig. 69). The dura ends at the lower border of the second sacral, the spinal cord at the lower border of the first lumbar vertebra. The spinal theca can therefore be punctured in the lower lumbar region without fear of injury to the cord.

The arachnoid is separated from the pia by a space of variable size called the *subarachnoid space*, which contains the cerebrospinal fluid. The pia invests the nervous substance very closely. The arachnoid, however, does not dip into the sulci or fissures (with the exception of the longitudinal sulcus), and invests the spinal cord quite loosely. There are also definite dilatations of the subarachnoid space called *cisternæ*. The cisterna magna [cerebello-medullaris] is found in the interval between the medulla and the under-surface of the cerebellum. The cisterna pontis lies on the ventral aspect of the pons and contains the basilar artery. The cisterna basalis [inter-peduncularis] is formed by the arachnoid bridging across the interval between the tips of the temporal lobes, and contains the circle of Willis [circulus arteriosus]. Prolongations of the subarachnoid space extend along the sheaths of the spinal and cranial nerves, particularly the optic (p. 125). The rest of the cerebrospinal fluid lies in the *ventricles of the brain*. The ventricles establish a connection with the extraventricular fluid through the foramen of Magendie [medial aperture] in the middle line in the inferior part of the roof of the fourth ventricle, and the foramina of Luschka [lateral apertures] at the extremities of the lateral recesses of this ventricle. As the arteries and veins enter and leave the brain substance they are surrounded by the perivascular spaces, which are continuous at one end with the sub-arachnoid space and at the other with the fine spaces which surround the nerve cells (see Fig. 69). The flow along these perivascular spaces (which

[1] Assoc. Res. nerv. ment. Dis. (Baltimore): *Cerebrospinal Fluid*, 1924; *Intracranial Pressure in Health and Disease*, 1929. Greenfield and Carmichael, *Cerebrospinal Fluid in Clinical Diagnosis*, London, 1925. Neel, *Cells and Proteins of Cerebrospinal Fluid*, London, 1939. Friedemann, Physiol. *Rev.*, 1942, *22*, 125.

correspond to the spaces containing interstitial fluid in other organs) is normally outwards to the subarachnoid space, and they serve to remove waste products resulting from cell activity.

Choroid Plexuses.—Some arteries (accompanied by a covering of pia) pass through the brain substance to reach the lining ependymal layer in the lateral, third and fourth ventricles. They then break up into complex capillary networks which project into the ventricular cavities and become lined by the now much folded ependymal cells. The blood vessels and their lining epithelium constitute the *choroid plexuses*. The epithelium becomes differentiated

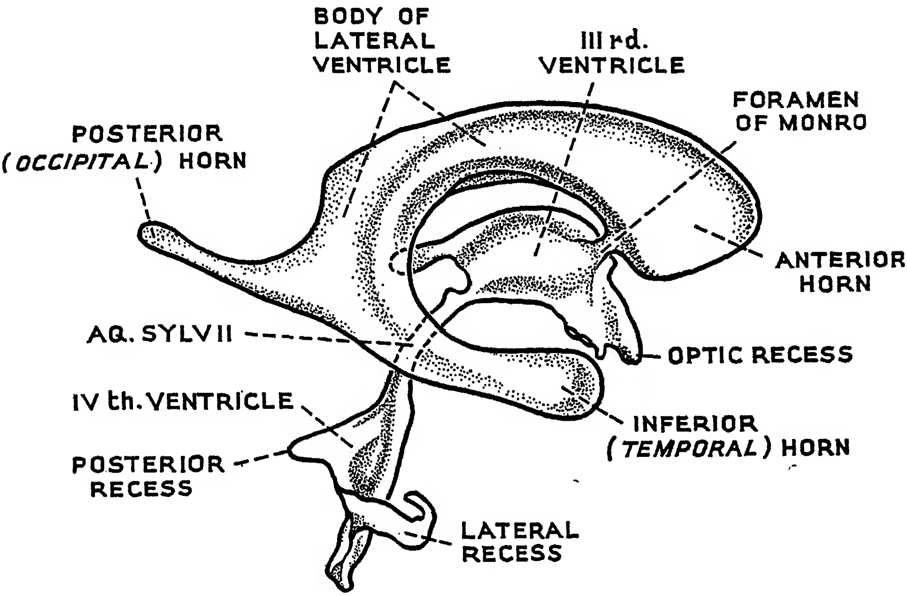


FIG. 64.—Lateral View of Cast of Human Ventricles. (After Retzius from Tilney and Riley.)

into cubical cells containing mitochondria, granules, and vacuoles, evidence that the cells are the seat of active metabolic processes.

COMPOSITION.—Cerebrospinal fluid is clear, colourless, and alkaline; the specific gravity is about 1005. It contains up to 5 lymphocytes per c.mm. In general its composition resembles that of protein-free plasma, but there are significant differences in the concentrations of the individual crystalloid constituents. The average values for its main constituents (in mg. per 100 c.c.) are as follows (the figures in brackets are those for plasma): Na^+, 334 (330); K^+, 10·6 (17); Ca^{++}, 5·3 (10·3); Cl′, 436 (365); HCO_3', 105 (150); HPO_4'', 1·8 (3·0); SO_4'', 0·6 (1·9); glucose, 70 (100). The *protein* content is minute, *i.e.* 0·02–0·035 g–% (plasma, 8·0 g–%), equally distributed between albumin and globulin. The intraventricular fluid is probably free from cells and protein; these may be added to the fluid in the subarachnoid space

by exudation from the meningeal blood vessels. The total *volume* of the cerebrospinal fluid in man is 100–150 c.c. When free escape is allowed to the exterior, the rate of formation is 20 c.c. per hour or about 500 c.c. per day.

Ventriculography.[1]—The anatomical condition of the cerebral ventricles can be studied clinically by injecting *air* into them. A small trephine hole is made in the skull and a fine needle is passed through a " silent area " of the brain into the lateral ventricle ; when it is entered, cerebrospinal fluid wells out. If the fluid is found to be under excessive pressure it is allowed

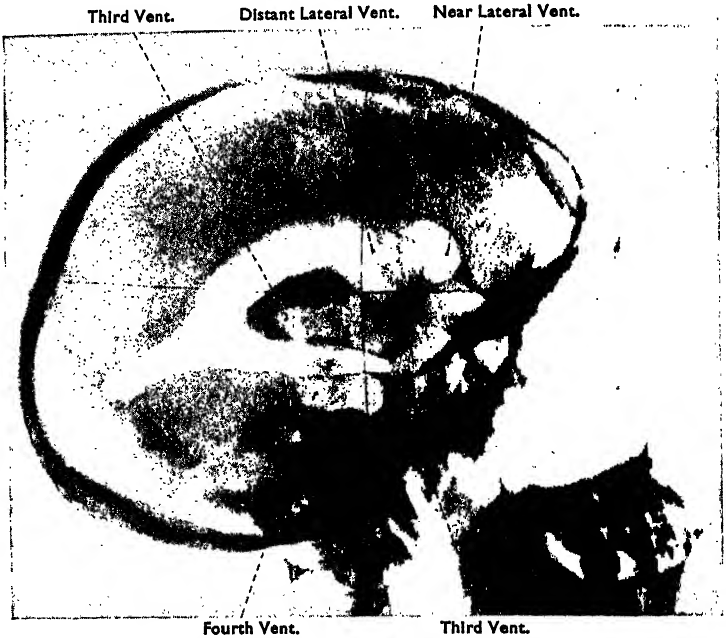


Fig. 65.—Normal Ventriculogram (Lateral View). (Harvey Jackson.)

to escape till the pressure has fallen to the normal level. Air is injected till it completely replaces all the fluid and finally fills the ventricles. A radiogram is then taken in the lateral and antero-posterior positions. Figs. 64 and 65 enable a comparison to be made between a cast of the human ventricles and a ventriculogram in a normal subject.

Ventriculography is of great value in determining the position, shape, and size of the ventricles and the patency of the various communications. The ventricles may be found to be excessively distended or incompletely filled, distorted, or displaced. Considerable help may be obtained in this way in the localization of cerebral tumours.

Formation of Cerebrospinal Fluid.—Site of Formation.—It is formed by the choroid plexuses, especially by the large plexuses which are found in the lateral ventricles. The cerebrospinal fluid passes from the lateral

[1] Loyal Davis, *Intracranial Tumours*, New York, 1933.

ventricles through the foramina of Monro [interventricular foramina] into the third ventricle, aqueductus Sylvii [cerebral aqueduct], and fourth ventricle, and out through the foramina of Luschka and of Magendie into the subarachnoid space, both cerebral and spinal (Fig. 64). The evidence is as follows.[1]

(i) If the interior of the lateral ventricle is exposed at operation in man, drops of clear fluid can be seen exuding from the surface of the choroid plexus.

(ii) Experimentally, if a catheter is introduced into the third ventricle, a continuous flow of fluid is obtained. On the other hand, if the aqueductus Sylvii is occluded, the lateral ventricles become greatly distended by the retained fluid, the condition being called *internal hydrocephalus*.

(iii) When one foramen of Monro is occluded, unilateral hydrocephalus (of the corresponding lateral ventricle) results. If, however, the corresponding choroid plexus is coincidentally extirpated the ventricle remains collapsed.

(iv) Obstruction of the foramina of Magendie and Luschka results in distension of the whole ventricular system.

Mechanism of Formation.—Cerebrospinal fluid may be regarded as a specialized form of interstitial fluid. One may consider whether its formation can be accounted for in the same way as interstitial fluid elsewhere,[2] *i.e.* whether it is due to the filtering force of the capillary blood pressure exceeding the opposing (suction) force of the plasma protein osmotic pressure. The following points should be noted: (i) the blood pressure in the capillaries in the choroid plexuses is unknown; (ii) while the hydrostatic pressure in the tissue spaces generally is zero (*i.e.* atmospheric pressure), the hydrostatic pressure of the cerebrospinal fluid in the ventricles is 5–10 mm. Hg. For filtration to occur in the choroid plexuses, the capillary pressure there must exceed the *sum* of the two opposing forces, *i.e.* the plasma protein osmotic pressure and the intraventricular cerebrospinal fluid pressure. It is impossible to say therefore whether adequate physical forces are available in the choroid plexuses to drive fluid out of the blood vessels into the ventricles. The capillaries in the plexuses, too, are not lined by a simple, thin endothelium, but are covered by a thick, highly differentiated and relatively impermeable, cubical epithelium. The crystalloid concentrations in plasma and cerebrospinal fluid (p. 119) differ in a manner and to a degree which cannot be accounted for by the known physical forces.

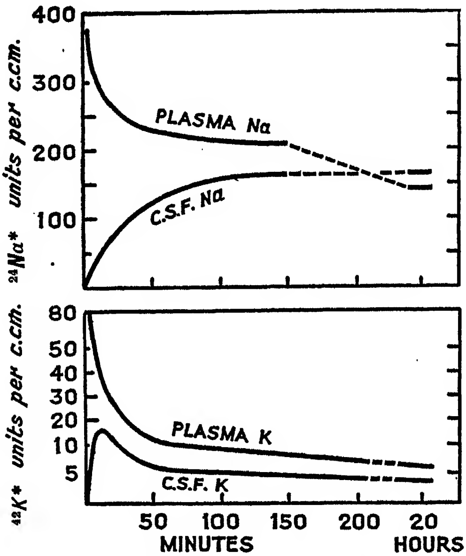
If cerebrospinal fluid were a simple filtrate its composition should follow rapidly and faithfully changes in the plasma. This point can be tested by injecting intravenously salts containing minute "tracer" doses of radioactive electrolytes and comparing their concentrations at intervals in the plasma and in *freshly formed* cerebrospinal fluid (collected from the cisterna magna). Representative results in such experiments after intravenous injection of salts containing radio-active sodium ($^{24}Na^*$), radio-active potassium ($^{42}K^*$), and radio-active phosphate ($^{32}PO_4'''^*$) are illustrated in Figs. 66, 67, and 68. These substances readily diffuse into the interstitial fluid and their concentrations there and in the plasma soon become equal. The figures, however, show that the *plasma and cerebrospinal fluid concentrations pursue an independent course* and even when finally some degree of

[1] Dandy and Blackfan, *Amer. J. Child. Dis.*, 1914, *8*, 406. Dandy, *Ann. Surg.*, 1919, *70*, 129.

[2] Read carefully pp. 17 *et seq.*

equilibrium is established there are marked differences in the concentrations in the two systems. The detailed findings are as follows:

(i) Plasma $^{24}Na^*$ falls rapidly during the first hour as equilibrium is established by diffusion into the interstitial fluid; the $^{24}Na^*$ level, however, rises more slowly in the cerebrospinal fluid than in the interstitial fluid elsewhere; when after many hours final equilibrium is attained cerebrospinal fluid $^{24}Na^*$ is found to *exceed* the plasma $^{24}Na^*$ level, a state of affairs that cannot be accounted for by simple diffusion through a "non-selective" membrane.



FIGS. 66 and 67.—Interchanges of Radio-active Na ($^{24}Na^*$) and K ($^{42}K^*$) between Plasma and Cerebrospinal Fluid (in dogs). (Greenberg *et al.*, *Amer. J. Physiol.*, 1943, *140*, 51.)

The ordinates represent the concentration of $^{24}Na^*$ (Fig. 66) and $^{42}K^*$ (Fig. 67) in arbitrary units in plasma (upper record) and cerebrospinal fluid (lower record) following an intravenous injection of a minute dose of salts containing these radio-active elements.

(ii) The behaviour of $^{42}K^*$ is quite different; the maximum level in the cerebrospinal fluid is reached in 10–20 minutes but is always *less* than in plasma; $^{42}K^*$ is progressively removed from the plasma by the muscles, other tissues, and red cells. The muscles ultimately retain over 90% of the injected element and the final muscle $^{42}K^*$ concentration may be four times as great as plasma $^{42}K^*$. [These results provide further evidence that K^+ ions *can* traverse muscle and other cell membranes (cf. p. 7).]

(iii) Radio-active P ($^{32}P^*$) (as phosphate) reaches its cerebrospinal fluid maximum in about 50 minutes and then declines, but always remains at a substantially lower level than in plasma. Plasma $^{32}P^*$ progressively decreases as it is incorporated in the *bones* or built up in the cells into organic esters.

These studies show that interchanges between plasma and cerebrospinal

fluid follow a different course from those between plasma and the interstitial fluids generally. The cells of the choroid plexuses thus maintain a characteristic electrolyte structure in the cerebrospinal fluid which differs from that of plasma ; in other words these cells *actively regulate* the transfer of ions and other crystalloids (cf. cell membranes, p. 7.) Cerebrospinal fluid formation is thus not a simple process of filtration and is therefore commonly labelled a process of "secretion," the word "secretion" drawing attention to the undetermined factors involved.

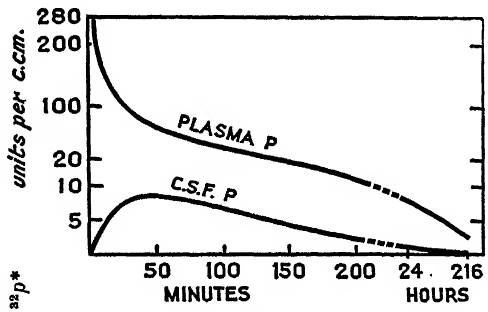


FIG. 68.—Interchange of Radio-active Phosphorus ($^{32}P^*$) (as PO_4''') between Plasma and Cerebrospinal Fluid (in dogs). (Greenberg *et al.*, *Amer J. Physiol.*, 1943, *140*, 54.)

The ordinate represents the concentration of $^{32}P^*$ (in arbitrary units) in plasma and cerebrospinal fluid following an intravenous injection of a salt containing a minute dose of $^{32}P^*$. (Note that the values on the ordinate are not uniformly spaced.)

It should be noted also that *bile pigment* does not penetrate into the cerebrospinal fluid (except in traces) even in deep jaundice. Similarly many *drugs* enter the cerebrospinal fluid from the blood only with difficulty.

Absorption.—(1) ROUTE.—Cerebrospinal fluid is absorbed mainly via the *arachnoid villi* into the dural sinuses and the spinal veins ; to a minor degree fluid may pass along the sheaths of the cranial nerves into the cervical lymphatics and also into the perivascular spaces. Roughly four-fifths of the fluid is absorbed via the cerebral arachnoid villi and most of the rest via the spinal villi. The arachnoid villi are small finger-like projections (lined by the usual flat epithelial cells of the subarachnoid spaces) which project into the venous sinuses as shown in Fig. 69. The Pacchionian bodies [arachnoideal granulation] are simply exceptionally large villi ; they are few in number and present only in the adult.

The route of absorption of the cerebrospinal fluid was demonstrated by Weed, who injected into the subarachnoid space of the living animal a solution of potassium ferrocyanide and iron ammonium citrate in isotonic saline, at a pressure of 15 cm. H_2O. The nervous system and meninges were removed and fixed in acid solution which precipitated the salts as Prussian blue. Histological examination showed that the dye passes through the arachnoid villi into the dural venous sinuses (and into the perivascular channels too if the animal is bled during the experiment). Colloids are also absorbed, but slowly.

(2) MECHANISM.—Absorption probably depends both on filtration and osmosis. In cats the cerebrospinal fluid pressure actually exceeds the pressure in the venous sinuses by 0·5–5 cm. H_2O, thus driving fluid into the veins (by *filtration*). More important probably is the 25 mm. Hg *osmotic pressure* of the plasma proteins, which draws fluid into the blood (in the same way that it causes the absorption of tissue fluid elsewhere, p. 15). As the cerebrospinal fluid pressure exceeds that in the veins, it might have been thought that the latter would be compressed against the skull bones and be obliterated. The firm attachments of the dural layer forming the walls of the veins is believed to prevent this happening.

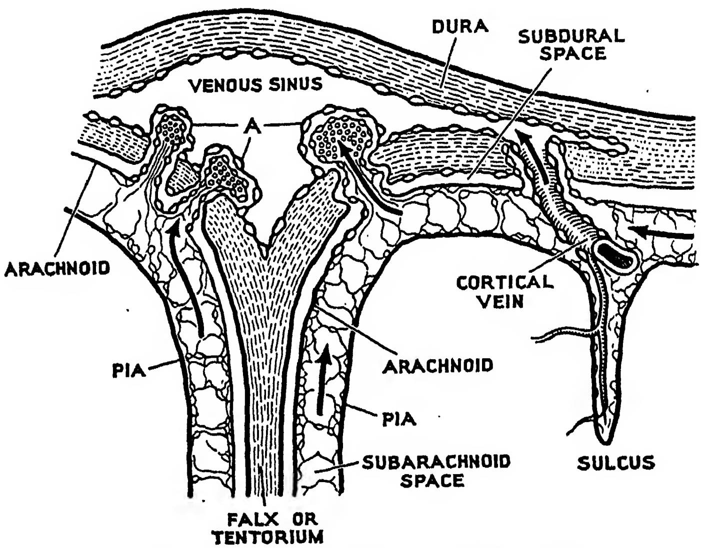


FIG. 69.—Diagram to show relations of Pia-arachnoid, Arachnoid Villi, and Cortical Veins to Dural Sinuses. Note the blood vessel coming from the brain substance surrounded by a continuation of the subarachnoid space. The arrows show the direction of flow of the cerebrospinal fluid. A = arachnoid villi. (Modified from Weed.) (H. Cushing, *Intracranial Physiology and Surgery*, 1926.)

Fluid Pressure.—The normal pressure of the fluid depends on a balance between its rate of secretion and of absorption. The value in man in the lateral recumbent position varies between 100–200 mm. H_2O. The pressure in the sitting position is 200 mm. H_2O higher than in the recumbent position. A rise of venous pressure such as follows coughing or crying hinders absorption and so raises the cerebrospinal fluid pressure. Compression of the internal jugular vein has a similar effect (*Queckenstedt's sign*) (Fig. 72).

Functions of Cerebrospinal Fluid.—(i) It serves as a fluid buffer. (ii) It acts as a reservoir to regulate the contents of the cranium; if the volume of the brain or of the blood increases, cerebrospinal fluid drains away; if the brain shrinks, more fluid is retained. (iii) It may serve to some extent as a medium for nutrient exchanges in the nervous system; in the main, however, the *brain carries out its metabolic exchanges directly with the blood.*

Effects of Cerebral Tumour and other "Space-Occupying Lesions."—If there is any increase in the intracranial contents, as by a tumour of the cerebrum, additional room is made by *expulsion* of some cerebrospinal fluid. Then the blood vessels are compressed and the gyri flattened, and gradually there is a general rise of pressure above the tentorium. This rise of pressure is transmitted to the prolongations of the subarachnoid space which surround the *optic nerves.* The first structures in the nerve to suffer from compression are the veins ; blood can still flow along the arteries and reach the optic disc, but its return is interfered with. In consequence, the minute vessels at the nerve head become engorged and swollen, and fluid exudes from them at this point of least resistance. The resulting appearance, on ophthalmoscopic examination, is termed *papillœdema.*

Pressure is then exerted on the posterior fossa of the skull. The cerebellum is gradually driven into and through the foramen magnum, fills this aperture like a cork, and thus impedes the escape of fluid into the spinal canal, whence about one-fifth of the total absorption of fluid normally occurs. In addition the fourth ventricle foramina (of Magendie and Luschka) are probably distorted and partially blocked. A vicious circle is thus established : as cerebrospinal fluid cannot escape from the ventricles and is not absorbed, hydrocephalus (*infra*) results. This further raises the intracranial pressure and wedges the cerebellum still more firmly into the foramen magnum ; the cerebral vessels are further compressed and finally, death results from medullary anæmia.

Hydrocephalus.[1]—By this term is meant a pathological accumulation of cerebrospinal fluid ; the hydrocephalus may be *internal* or *external*, according as to whether the excess fluid is in the ventricular system or in the subarachnoid space. Theoretically hydrocephalus might be due to :

(i) *Oversecretion* of fluid, the rate of formation exceeding the rate of absorption. Oversecretion has been alleged to result from (*a*) extensive hypertrophy of the choroid plexuses, or (*b*) a rise of the capillary blood pressure in the plexuses owing to thrombosis of the vein of Galen which drains this region. Clinically oversecretion is a negligibly rare cause.

(ii) *Obstruction* to the outflow of the fluid. The obstruction may be :

(*a*) *Intraventricular*, blocking the foramen of Monro, the cerebral aqueduct, or the fourth ventricle foramina. The fluid in the occluded ventricles cannot escape and cannot undergo absorption locally ; its volume progressively increases because of continued formation of fluid by the choroid plexuses.

(*b*) *Extraventricular* : (*a*) preventing the free flow of the fluid throughout the subarachnoid space and so diminishing the total surface of arachnoid villi available for absorption. Thus a block at the foramen magnum prevents the fluid from entering the spinal arachnoid and thus cuts off one-fifth of the absorbing surface ; a block at the tentorial opening [cisterna ambiens] prevents the fluid from passing from the posterior fossa of the skull into the supratentorial subarachnoid space where most of the absorption normally occurs ; (β) inflammatory changes in the leptomeninges may occlude the arachnoid villi ; (γ) thrombosis of the dural sinuses prevents the escape of the fluid from the subarachnoid space into the veins.

[1] Dorothy S. Russell, *Pathology of Hydrocephalus*, British Med. Res., Council Sp Rep. No. 265, London, 1949.

Experimental Hydrocephalus.—(i) The injection of foreign particles (*e.g.* lamp black, trypan blue, or thorotrast) into the cisterna magna stimulates the arachnoid cells (*i.e.* those lining the subarachnoid space). They become rounded, proliferate rapidly, and display phagocytic properties, ingesting the foreign material like macrophages elsewhere (cf. p. 187). Secondarily the leptomeninges become matted together and the arachnoid villi are presumably blocked, thus interfering with the flow and absorption of the cerebrospinal fluid.

(ii) An alternative procedure consists in making repeated intracisternal injections of the animal's own blood; the blood is laked, the released hæmoglobin is taken up by the macrophages locally and bile pigment is formed which colours the fluid. Meningeal reactions occur as in (i) and likewise cause hydrocephalus.

Clinical Hydrocephalus.—The causes are numerous: (i) *congenital malformations* within the ventricular system at the narrow points (usually the cerebral aqueduct); developmental abnormalities of the base of the skull leading to blocking of the foramen magnum; malformation of the brain in the neighbourhood of the foramen magnum which becomes occluded, *e.g.* by a caudal prolongation of the cerebellum; (ii) *infections* which may block the cerebral aqueduct, the fourth ventricle foramina, or the tentorial opening, or more commonly destroy the absorbing mechanisms in the subarachnoid space; (iii) *tumours* may deform and obstruct the intraventricular system and set up a vicious circle by raising intracranial pressure as described on p. 125; (iv) occasionally dural sinus thrombosis is the causal agent.

Communicating and Non-Communicating Hydrocephalus.—The site of the obstruction can often be determined by injecting phenol sulphonephthalein into the lateral ventricle. Normally it should appear in fluid obtained by lumbar puncture in 2–3 minutes, and in the urine in 10–12 minutes. If it does not appear in the spinal fluid or only after a long delay the hydrocephalus is *non-communicating*, *i.e.* the site of the block is not distal to the fourth ventricle foramina. If it appears in the spinal fluid normally but in the urine after an undue delay, the hydrocephalus is *communicating*, *i.e.* the block is in the meninges distal to the fourth ventricle foramina, in the arachnoid villi, or in the dural venous sinuses.

Effects of Injection of Saline Solutions.[1]—(1) The intravenous injection of large amounts of *Ringer's solution* causes a temporary rise of cerebrospinal and venous pressure; the condition soon returns to normal (cf. p. 59). The rise of venous and capillary pressure and the dilution of the plasma proteins promote cerebrospinal fluid formation and hamper its absorption (as with tissue fluid elsewhere).

(2) The injection of *hypotonic* saline (or distilled water) causes a marked and *prolonged rise* of cerebrospinal fluid pressure (Fig. 70, A). There is an associated increase in *venous pressure,* but this is transient and so cannot be an important causal factor. The decreased crystalloid osmotic pressure of the plasma causes water to pass out of the plasma into the interstitial spaces where the osmotic pressure is higher. [Salt diffuses the other way; but the results show that the movement of the water is so much more rapid and extensive that *initially* it alone need be considered.] Similarly water flows from the plasma into the cerebrospinal fluid; later there is a flow of water

[1] Weed and Hughson, *Amer. J. Physiol.*, 1921, *58*, 53.

from the extracellular compartment into the cells including the brain cells which become *swollen* (cf. p. 64). There is thus a rise of cerebrospinal fluid pressure, a rise of intracranial pressure, and *swelling of the brain.*[1]

(3) If *hypertonic* saline (*e.g.* 30% NaCl) is injected intravenously the cerebrospinal fluid pressure (after a sharp, brief rise) *falls* profoundly for a period of 2–4 hours, recovery usually occurring within 7 hours. This effect is independent of the arterial pressure, which shows variable changes, and of the venous pressure, which rises initially and then returns to about its original level (Fig. 70, B). There is marked *shrinking of the brain substance:*

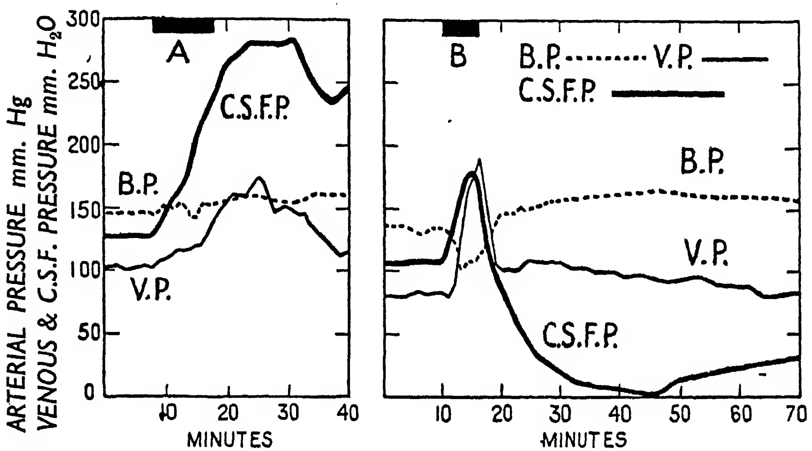


FIG. 70.—Effects of Hypotonic and Hypertonic Saline on Cerebrospinal Fluid Pressure, Venous Pressure, and Arterial Blood Pressure. (Experiments on cats.) (Weed and Hughson, *Amer. J. Physiol.*, 1921.)

Ordinate: C.S.F. pressure (C.S.F.P.) and venous pressure (V.P.) in mm. water; arterial blood pressure (B.P.) in mm. Hg.
A: inject 50 c.c. of distilled water intravenously. There is a trivial rise of arterial pressure, a rise of venous pressure and a marked sustained rise of C.S.F. pressure.
B: inject 12 c.c. of a 30% NaCl solution intravenously. There is a slight rise of arterial blood pressure after an initial fall, and a transient rise of venous pressure. After an initial rise, C.S.F. pressure shows a very marked and sustained fall.

the convolutions become smaller and the sulci wider and deeper. The injection of 1 g. of NaCl per kg. body weight lowers the cerebrospinal fluid pressure by 8 cm. H_2O. If 10–20 c.c. of a 30% solution of NaCl are introduced into the duodenum or rectum of a cat the cerebrospinal fluid pressure may fall by 16 cm. H_2O and not be restored to normal for 17–48 hours; if the fluid is introduced into the stomach the results are less certain; saturated solutions of sodium sulphate give similar results, but are slower and less marked in their action. The general effects of an injection of hypertonic saline on the volume and composition of the extracellular fluid (which includes the plasma) and the intracellular fluid are shown in Fig. 71.

Hypertonic saline acts by raising the crystalloid osmotic pressure of the plasma. Water consequently flows from the interstitial fluid into the plasma and secondly *from the cells into the interstitial fluid.* The *brain* (like cells

[1] Similar changes occur in *water intoxication* (p. 58) and may largely account for the cerebral symptoms.

elsewhere) therefore shrinks, and cerebrospinal fluid is withdrawn osmotically into the cerebral blood vessels generally. In addition there is a *reversal of cerebrospinal fluid flow*, i.e. *it passes backwards through the choroid plexuses from the ventricles into the blood vessels.* Using Weed's Prussian blue method, (p. 123), Foley [1] introduced the ferrocyanide and iron-ammonium citrate solution into the cisterna magna. Normally, even if the fluid is injected under a pressure of 40 cm. H_2O, the dye does not penetrate beyond the aqueductus. It was found, however, that after the intravenous injection of hypertonic saline, the dye passed freely into the perivascular spaces and into the blood vessels which they surround, into the vessels which course through the subarachnoid space, into the ventricular system and through the aqueductus Sylvii to the extremities of the frontal and occipital lobes. If a catheter connected to a manometer is introduced into the cerebral aqueduct so as to obstruct it, it is found that the intravenous injection of hypertonic

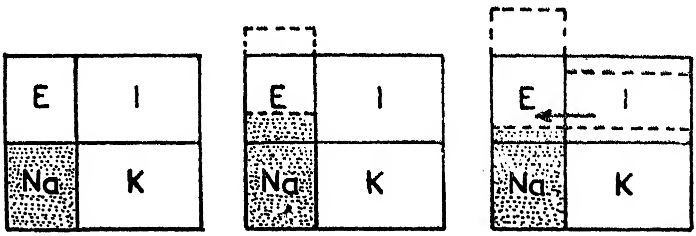


FIG. 71.—Effect of Hypertonic Saline on Distribution and Composition of Body Fluids.

Total areas of E, I represent volume of extracellular and intracellular fluid. Areas Na, K represent the concentration of these ions in extracellular and intracellular fluid.
Left-hand box: Control values.
Middle box: Effects of injection of hypertonic saline. The volume of extracellular fluid (E) is increased and its Na^+ concentration is raised. [The flow of interstitial and cerebrospinal fluid into the blood is not depicted.]
Right-hand box: The hypertonicity of the extracellular fluid withdraws fluid from the cells. The intracellular fluid volume (I) is decreased and its K^+ concentration is raised.

saline causes an immediate and profound *fall of intraventricular pressure, although the aqueduct is occluded* and the normal path of escape for the cerebrospinal fluid is not available. If Weed's fluid is introduced through the catheter directly into the ventricles, and the experiment is repeated, the dye is found to pass *through the cells of the choroid plexus into the contained capillaries.* As the intraventricular pressure changes precede those in the subarachnoid space, it is obvious that fluid is being *absorbed* by the plasma in the choroid plexuses (owing to differences of crystalloid osmotic pressure), thus setting up a further current of fluid from the subarachnoid space [cisterna magna] into the ventricles.

CLINICAL APPLICATIONS.—The raised intracranial pressure of cerebral tumour is temporarily lowered by intravenous injection of hypertonic solutions. The solutions employed are: 10% saline (50 c.c.), 25% glucose (100 c.c.), or 50% sucrose (100 c.c.). Consciousness may be restored, headache is relieved, and the swelling of the optic disc diminished. Intracranial

[1] Foley, *Arch. Neurol. Psychiat.*, 1921, *5*, 744; *Arch. Surg.*, 1923, *6*, 587.

operations are made easier and safer as excessive bulging of brain substance is prevented. The beneficial effect of hypertonic solutions is *transient.* The hypertonic solution can withdraw c.s.f. into the blood only while the crystalloid o.p. of the blood exceeds that of the interstitial fluid (including c.s.f.). Gradually, however, the *salt diffuses out* of the plasma into the interstitial fluid; the water follows it; and finally the plasma and interstitial fluids come into osmotic equilibrium, both being slightly hypertonic and both being *increased in volume.* This stage is harmful if it were not compensated for by renal activity. Glucose is oxidized in the body leaving a condition of simple water excess which is readily dealt with by the kidney. Sucrose is rapidly excreted in the urine. With hypertonic saline the excretion of the injected water must, however, await the relatively slow renal excretion of the salt.

Cases of internal hydrocephalus benefit temporarily from this treatment to a marked degree owing to the reversal of flow of the cerebrospinal fluid and the absorption of the accumulated intraventricular fluid into the choroid plexuses and elsewhere (p. 128).

Concentrated human serum (containing, *e.g.* 30% of protein) may also be used clinically.[1] The resulting rise of plasma *protein* osmotic pressure (p. 15) leads to withdrawal of cerebrospinal fluid into the blood vessels, producing a fall of intracranial pressure which may last for 18–24 hours; the crystalloids of the concentrated plasma also produce their characteristic effect for the first few hours after injection.

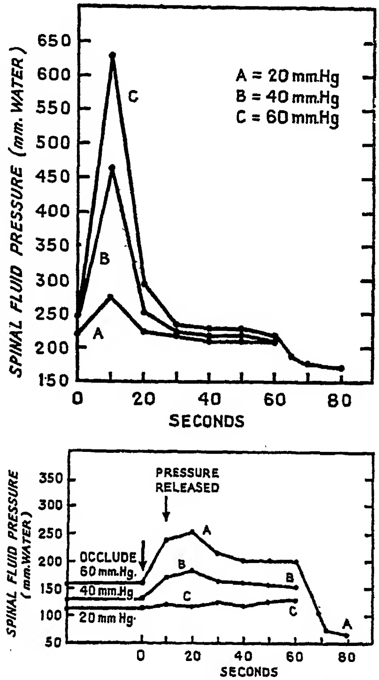


FIG. 72.—Effect of Raised Jugular Venous Pressure on Lumbar Cerebrospinal Fluid Pressure in Normal Subject and in Patient with Partial Spinal Block. (Turner and Byrne, *Yale J. Biol. Med.*, 1940, *12*, 739).

Upper record. Normal. A blood pressure cuff is fixed round the neck and the bag is inflated to pressures of 20 mm. Hg (A), 40 mm. Hg (B), 60 mm. Hg (C) at the beginning of the record (point 0 on abscissa), maintained for 10 sec. and then released. Note the rise in c.s.f. pressure in each case to about the level of the occlusion pressure. On releasing the cuff pressure the c.s.f. pressure rapidly falls. *Lower record.* Case with partial spinal block. The *initial* c.s.f. pressure is lower. Occluding the jugular vein for 10 sec. (between the arrows) produces a *much smaller rise* of c.s.f. pressure.

LOCULATION SYNDROME.—The spinal subarachnoid space may be obstructed by injecting paraffin wax. Remarkable changes occur in the fluid below the level of the block; the protein content is greatly increased and the colour is now yellow. The exact cause of these changes is not known. Possibly the spinal veins which run longitudinally upwards are compressed; the venous and capillary blood pressures rise, and consequently blood corpuscles

[1] Hughes, Mudd and Strecker, *Arch. Neurol. Psychiat.*, 1938, *99*, 1276.

and plasma containing much protein exude. The hæmoglobin disintegrates and gives rise to a yellow pigment and the carotene of plasma also contributes to the colour. Clinically, chronic meningitis or tumours of the cord and its envelopes produce identical changes in the cerebrospinal fluid—the *Froin* or *loculation syndrome.* The protein is increased to 0·5 or 1%, or even up to 4%; albumin and globulin are present in the same proportions as in blood; the fluid coagulates spontaneously, owing to the presence of fibrinogen; the yellow coloration is termed *xanthochromia.* The cell count is normal, except in syphilitic cases. The only essential feature of the syndrome is the raised protein content above 0·5%. (In inflammatory lesions of the meninges the increased permeability of the blood vessels also permits the exudation of a fluid rich in protein.)

When the spinal subarachnoid space is blocked, compression of the internal jugular veins in the neck raises the cerebrospinal fluid pressure above the level of the block to the normal extent; below the level of the block the pressure of the fluid is raised very little or not at all (Fig. 72, lower record).

Lumbar Puncture.—This is carried out by introducing a needle into the subarachnoid space below the termination of the spinal cord, usually between the spines of the fourth and fifth lumbar vertebræ.

INDICATIONS:

(1) For diagnostic purposes.
(2) To relieve raised intracranial pressure, *e.g.* in meningitis or uræmia.
(3) For the introduction of therapeutic agents.
(4) To produce spinal anæsthesia.

USE OF LIPIODOL.—Lipiodol is a very heavy liquid, opaque to X-rays, which can be injected into the cisterna magna. Normally it drops by reason of its weight to the bottom of the spinal theca. If the subarachnoid space is obstructed by tumour or by adhesions, the lipiodol is held up at that point and its position can be demonstrated by X-rays.

Pathology of Cerebrospinal Fluid.—The significance of certain changes in the composition of the cerebrospinal fluid will now be considered.

PROTEIN.—The amount of protein is estimated by boiling with trichloracetic acid and comparing the turbidity produced with that of fluids containing known concentrations of protein.

If *globulin* is present in excessive amounts, the *Nonne-Apelt reaction* is obtained. On adding 1 c.c. of cerebrospinal fluid to 1 c.c. of saturated ammonium sulphate solution a grey ring of precipitate appears at the junction of the two fluids.

The amount of protein is increased in all meningeal affections and is the expression of increased permeability of the inflamed blood vessels. A positive test for globulin when the total protein is 0·1% or less is strongly suggestive of a syphilitic affection.

CELLS.—An increase in the number of cells in the fluid is suggestive of inflammation.

(i) *Lymphocytosis* usually occurs in acute infections, particularly in tubercular meningitis, lethargic encephalitis, acute anterior poliomyelitis (except in the pre-paralytic stage), and in all syphilitic affections of the nervous system. In the last group the number of lymphocytes depends on the degree

of meningeal involvement, and the number of cells may vary between 10 and 500 per c.mm.

(ii) *Leucocytosis* (polymorphonuclear) occurs in all pyogenic infections, *e.g.* meningococcal or pneumococcal, and some cases of tubercular meningitis. Large *mononuclear* cells appear in cases of tumours invading the meninges.

GLUCOSE and UREA follow the level of these substances in the blood. The glucose concentration is high in diabetes, and the urea level is usually raised in uræmia.

The general changes in the cerebrospinal fluid in *meningitis* are: increased pressure, increased cell count, increased protein concentration, and decreased concentration of glucose and chloride. Acute *purulent* meningitis is mainly due to meningococcus, pneumococcus, streptococcus, and staphylococcus; the cells are mainly polymorphs. In *tubercular* meningitis the fluid changes are generally less marked; the predominant cell is the lymphocyte.

The following Tables (after Symonds) show the changes in the cerebrospinal fluid in some important conditions:

	Normal Cerebro-Spinal Fluid.	Acute Purulent Meningitis.	Tubercular Meningitis.
Colour . .	Clear and colourless.	Turbid or purulent. Clot on standing.	Clear, thread-like coagulum on standing.
Protein . .	Average, 0·02%.	Increased to 0·1–0·5%.	Increased to 0·06–0·3%.
Sugar . .	0·08%.	Diminished or absent.	Diminished, but not constantly.
Cl′ (expressed as NaCl) .	0·77%.	Reduced to 0·65 or 0·6%.	Reduced, often below 0·6%.
Cells per c.mm.	1–5 lymphocytes per c.mm.	Increased. 1000–10,000 of polymorphs.	Increased. 50–500. Usually 75% lymphocytes.
Pressure in mm. water in lateral recumbent position.	100–200.	Increased. 200–500 or more.	Increased.
		Causative organism present.	Tubercle bacilli may be present.

	General Paralysis.	Tabes.	Meningo-Vascular Syphilis.
Protein . .	0·05–0·1%. Globulin reaction +.	0·03–0·08%. Globulin reaction +.	0·03–0·08%. Globulin reaction +.
Cells . . .	Up to 400 per c.mm. All lymphocytes.	10–80 per c.mm. or more. All lymphocytes.	10–80 per c.mm. All lymphocytes.
Wassermann reaction—			
Blood . .	Early, 75% +. Late, 100% +.	70% +.	80% +.
Cerebrospinal fluid .	All stages, 99% +.	70% +.	50% +.
Benzoin curve.	Paretic.	Luetic.	Luetic.

COLLOIDAL BENZOIN REACTION.—Normal cerebrospinal fluid in any dilution added to a colloidal benzoin solution has no effect. In syphilis of the nervous system, incomplete flocculation (recorded as 1) or complete

II

THE BLOOD [1]

The blood consists of plasma and corpuscles (red cells, white cells, and platelets). The following aspects of the physiology of the blood have already been discussed :

(i) Composition of plasma and red cells (pp. 5, 7).
(ii) Interchanges between plasma and interstitial fluid (p. 17).
(iii) Total blood volume, plasma volume, red cell volume : their determination (p. 9) and regulation (p. 80) in health and disease ; hæmatocrit value (p. 10).
(iv) H^+ ion concentration of the blood in health and disease (p. 87).
(v) Rôle of the kidney in the regulation of the composition and volume of the plasma (p. 21).

Other aspects of the physiology of the blood are dealt with below.

THE PLASMA PROTEINS [2]

The total plasma protein concentration is 6·4–8·5 g. per 100 c.c. Two principal groups of plasma proteins are conventionally recognized : *serum albumin and serum globulin.* The globulin fraction can be further subdivided into α-, β-, and γ- globulin and fibrinogen. A number of proteins with specific physiological functions have been partially isolated from the globulin fraction by such methods as electrophoresis. Among these proteins are prothrombin, plasma thromboplastin, isohæmagglutinins, hypertensinogen, immune globulins, and anterior pituitary hormones.

Properties of Proteins.—Only those properties which need be known to understand the methods of separating the proteins of the plasma, and their *physiological* properties, will be considered here.

(1) Precipitation by Salts.—Different proteins are precipitated from solution by addition of different concentrations of salts. Thus albumin is precipitated by saturation with $(NH_4)_2SO_4$, globulin by half-saturation with $(NH_4)_2SO_4$. This method of separation gives a normal serum albumin/globulin ratio of 1·7 (*e.g.* albumin 4·5, globulin 2·7 g–%).

(2) Sedimentation in Ultracentrifuge.—The different proteins sediment at different rates when solutions of them are spun at very high speeds in the Svedberg ultracentrifuge : separation can thus be effected.

[1] See Whitby and Britton, *Diseases of Blood,* 6th edn., London, 1950. Wintrobe, *Clinical Hæmatology,* Phila., 2nd edn., 1946. Downey, *Handbook of Hæmatology,* N.Y., 1938 (4 vols.).

[2] Marrack and Hoch, *J. clin. Path.,* 1949, *2,* 161. Cohn and coworkers, *J. clin. Investig.,* 1944, *23,* 417 and 21 succeeding papers ; Cohn, *Blood,* 1946, *1,* 3. Gutman (Plasma Proteins in Disease), *Advances in Protein Chemistry,* 1948, *4,* 155.

(3) Isoelectric Point.—Proteins can behave either as bases or as acids owing to the fact that their constituent amino-acids (p. 876) contain both *basic* amino (*i.e.* NH_2) groupings and *acidic* carboxyl (*i.e.* COOH) groupings. A protein can be diagrammatically represented thus: $\begin{matrix} NH_2 \\ COOH \end{matrix}\!\!>\!R$. In an alkaline solution a protein behaves as an acid (by virtue of its COOH groups) and forms a salt ($NH_2{-}R{-}COONa$); in an acid solution it behaves as a base (by virtue of its NH_2 groups) combining with acids, *e.g.* HCl, to form $NH_3Cl{-}R{-}COOH$.

In solution a protein ionizes slightly. In *alkaline* solution the reaction can be represented thus:

$$R\begin{matrix} \diagup NH_2 \\ \diagdown COONa \end{matrix} \rightleftharpoons R\begin{matrix} \diagup NH_2 \\ \diagdown COO' \end{matrix} + Na^+$$

i.e. the protein ion is *negatively* charged (it is an *anion*).

In *acid* solution the reaction is:

$$R\begin{matrix} \diagup NH_3Cl \\ \diagdown COOH \end{matrix} \rightleftharpoons R\begin{matrix} \diagup NH_3^+ \\ \diagdown COOH \end{matrix} + Cl'$$

i.e. the protein ion is *positively* charged (it is a *cation*).

At a certain *p*H (specific for each protein) a protein behaves neither as an acid nor as a base; it is un-ionized and does not move in an electric field. This *p*H is called the *isoelectric point*; at the isoelectric point the solubility of the protein is at its minimum and it tends to precipitate out. The isoelectric point of serum albumin is *p*H 4·7, that of the serum globulins is *d*H 5·4–7·0.

(4) Buffer Action.—As the *p*H of plasma is 7·4, *i.e.* well to the alkaline side of the isoelectric point of the plasma proteins, the latter behave as acids and are bound with base (almost entirely Na^+). Such molecules as are ionized give rise to protein anions. The reaction can be represented thus:

$$NaPr \rightleftharpoons Na^+ + Pr'.$$

The amount of base (Na^+) which is combined with the 70 g. of protein in one litre of plasma is 16 milliequivalents (cf. p. 12) (*i.e.* $16 \times 0{\cdot}023 = 0{\cdot}4$ g. of Na^+ per litre).

The base (B) which is combined with the plasma proteins can be attacked by H_2CO_3 to form $BHCO_3$ (actually $NaHCO_3$); this reaction helps to preserve the reaction of the plasma and aids in the carriage of CO_2. The plasma proteins are thus buffer substances (p. 91). Direct observation shows, however, that at the *p*H of the blood (which is far removed from their isoelectric point) the plasma proteins retain a firm hold on their base and so actually play little part either as buffers or as aids to CO_2 carriage; they are probably responsible for *less than one-tenth* of the total buffer action of blood.

(5) Molecular Weight.—The molecular weight of serum albumin is 69,000, that of γ-globulin 156,000 and that of fibrinogen about 500,000; these molecular weights are related to molecular size. Larger molecules pass less readily than smaller molecules through the capillary wall. The

normal capillary endothelium is almost impermeable to all the plasma proteins; but when its permeability is increased (*e.g.* in disease), albumin (having the smallest molecular weight) is the first protein to pass through.

(6) OSMOTIC PRESSURE.—See p. 15.

(7) MOLECULAR SHAPE AND DIMENSIONS.—Expressed in Ångström units (Å.U.) (Å.U.$=10^{-8}$ cm.) the dimensions are:

albumin, 150×38 Å.U.;
γ-globulin, 320×36 Å.U.;
fibrinogen, 900×33 Å.U.;

(cf. Na^+, $1{\cdot}9 \times 1{\cdot}9$; Cl', $3{\cdot}6 \times 3{\cdot}6$; glucose, $9{\cdot}5 \times 6{\cdot}5$ Å.U.). The shape of the molecule determines the viscosity of its solutions (*infra*).

(8) VISCOSITY.—The resistance to the flow of fluid (at constant velocity) through a capillary (of constant bore) depends almost entirely on the viscosity of the fluid. The viscosity of the blood is thus a factor in maintaining the peripheral resistance and thereby, the arterial blood pressure (p. 303). The viscosity of a protein solution depends far more on the *shape* of the protein molecule than on its *size*; the less symmetrical the molecule the greater is its viscosity. For this reason the following solutions have equal viscosities: 25% albumin, 15% γ-globulin, 2% fibrinogen; each of these solutions has a viscosity equal to that of twice concentrated plasma (*i.e.* a plasma with 15 g. of "mixed" proteins per 100 c.c.). Surprisingly enough the viscosity of whole blood (*i.e.* plasma plus suspended corpuscles) is also only that of twice concentrated plasma; this means that the corpuscles and the plasma contribute equally to the total viscosity of the blood.

(9) ELECTROPHORETIC MOBILITY.—As already explained, in acid solution proteins form cations and in alkaline solutions they form anions. If the force of an external electrical field is applied to protein molecules dissolved in a suitable solvent, they are caused to move; in acid solutions the protein ion (being positively charged) moves towards the cathode; in alkaline solutions the protein ion (being negatively charged) moves towards the anode. The rate of movement or *mobility* (μ) is expressed in cm. per sec. when the protein is in an electrical field of one volt per cm. The mobility varies with the protein (also with *p*H, viscosity of the solvent, and the nature and concentration of the dissolved salts). Each protein moves with a characteristic mobility; therefore when a solution of mixed proteins (in a tube) is placed in an electrical field with the electrical poles at the ends, the contained proteins gradually separate out like the runners in a long race who, having started together in a row, end up at the winning post as a long string. The greater the difference in mobility between the molecules, the more complete is the separation which can be effected by this method. Using a suitable technique the proteins can be photographed as they are "strung out" along the tube giving the kind of picture shown in Fig. 73; these pictures are called *electrophoretic patterns* or *diagrams*. The area under each peak is a measure of the concentration of the particular component. The principal peaks (in *descending* order of mobility) are albumin (highest mobility); α_1-, α_2-globulin; β-globulin; fibrinogen; γ-globulin (lowest mobility).

The concentrations of the protein fractions thus separated out (expressed as *percentages of the total plasma protein content*) are: albumin 55%; α-globulin

$\alpha_1+\alpha_2$) 13%; β-globulin 14%; fibrinogen 7% and γ-globulin 11%. The albumin/globulin ratio obtained by electrophoretic analysis is thus 55/45=1·2, an average result which is two-thirds that obtained by classical chemical fractionation, *i.e.* 1·2 compared with 1·7 (pp. 15, 133).

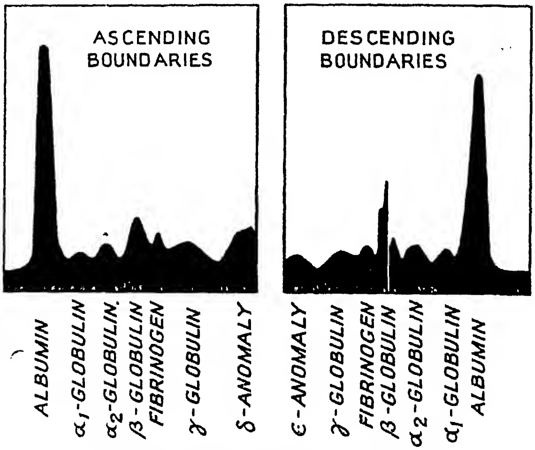


FIG. 73.—Electrophoretic Diagram of Normal Human Plasma. (Cohn *et al.*, *J. clin. Investig.*, 1944, *23*, 420.)

The area of each deflection is proportional to the concentration of the individual protein in the plasma. (Cf. Fig. 550)

Six main functional protein fractions have been obtained by electrophoresis of plasma; their components are shown below.

Fraction

I. Fibrinogen (+the globulin called plasma thromboplastin which promotes clotting in hæmophiliacs.)

mainly β and γ globulins:
II. Immune globulins (=γ-globulin).
III. (1) Isohœmagglutinins (=β-, γ-globulins).
(2) Prothrombin, fibrinolysin, plasma thromboplastin (=α-β-, γ-globulins).

IV. α- and β-globulins, hypertensinogen, alkaline phosphatase.

V. Albumin.

VI. Proteins in mother liquor (albumin and α-globulin); also follicle-stimulating hormone of anterior pituitary.

Many of the proteins mentioned above have been isolated in a high degree of purity, *e.g.* *albumin* (in a form suitable for intravenous injection clinically); *isohœmagglutinins* (in 16 times the concentration found in pooled plasma); *immune globulins* (active against diphtheria, influenza virus, measles, mumps, typhoid bacillus) with 15–30 times the immune potency of pooled plasma; proteins concerned in blood *clotting* (p. 139).

Relation of Diet to Plasma Proteins.[1]—This is best studied in the *standard plasma-depleted dog*, as described by Whipple. Whole blood is

[1] Madden and Whipple, *Physiol. Rev.* 1940, *20*, 194. Whipple, *Amer. J. med. Sci.*, 1942, *203*, 477.

withdrawn and the corpuscles reinjected suspended in Ringer-Locke's solution (*i.e.* a protein-free fluid); the procedure is known as *plasmapheresis* ("plasma skimming"), and if repeated daily leads to a progressive diminution in the concentration of plasma protein, as the rate of protein withdrawal exceeds the rate of regeneration. Depletion is continued for some weeks after the plasma protein concentration has fallen to 4%, in order to exhaust the protein reserves. Thereafter, on a standard diet, the rate of plasma protein formation is constant.

The results show, as might be expected, that plasma proteins are normally formed from food proteins, but that in protein starvation they may be formed from tissue protein. The efficacy of a food protein depends on the degree of its chemical resemblance in amino-acid pattern to the plasma protein which it is going to form; very naturally plasma proteins are the most efficient raw materials.[1] Plasma proteins can also be satisfactorily synthesized from amino-acids if the ten essential ones are present, *i.e.* arginine, valine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, threonine, tryptophane. (Specific experiments have shown that methionine is essential for long-continued plasma protein formation and also for adequate body nutrition; cystine can be substituted for short periods only) (cf. p. 881). As albumin and globulin have quite distinct amino-acid patterns, some proteins (*e.g.* those from muscle and viscera) favour albumin formation, while others (*e.g.* plant and grain proteins) favour globulin formation. The presence of *infection* depresses protein regeneration.

Origin of Plasma Proteins.—Serum *albumin* and the proteins concerned in blood *clotting* (fibrinogen, prothrombin) are probably formed exclusively by the liver. In disease of the liver (*e.g.* acute hepatitis, cirrhosis (Figs. 544 and 549)) the concentration of these constituents in the plasma falls markedly. Experiments with "labelled" plasma proteins suggest that they are completely destroyed and replaced every two weeks (p. 879).

The serum *globulins* are probably formed widely in the body, perhaps by the reticulo-endothelial system and the lymphoid nodules (p. 219); as the immune bodies are an important constituent of the serum globulin fraction, globulin formation is concerned with *resistance to infections*. Thus after three successive injections of typhoid vaccine the serum globulin rose from 2% ultimately reaching a level of 4%; (the fibrinogen level also rose (cf. p. 173)). The immune bodies are mainly in the γ-globulin fraction.

When the liver cells are extensively damaged the serum globulin fraction frequently rises, especially the γ-globulin, although the serum albumin falls; the consequent abnormal plasma protein pattern is probably responsible for the flocculation tests of liver function (p. 832).

After *hæmorrhage*, fibrinogen, serum globulin, and serum albumin are regenerated in that order, complete restoration being effected in a few days (p. 84).

In *nephrosis* (p. 75) serum albumin may be lost in the urine at the rate of 25 g. daily for several months before its concentration in the plasma is

[1] During *fasting* only 2–8 g. of plasma protein are formed weekly from tissue constituents. If *whole plasma* (fresh or dried) is given by mouth, every 3 g. of ingested protein produces 1 g. of new plasma protein; the potency ratio is thus 3 : 1. The proteins of grains, potato, kidney, or liver have a potency of 5 : 1; red cells, heart, or spleen, 10 : 1.

finally lowered—further evidence that the body can form this substance when necessary on a large scale.

In *infancy*, low total plasma protein concentrations are found (*e.g.* 5·1–5·5%) owing to the low albumin content. Albumin and globulin decrease in the first 6 months of *pregnancy*, but the fibrinogen increases.

The relationship between plasma proteins and tissue proteins is probably an intimate one. Whipple suggests that the proteins of the cells can be divided into three categories: (i) *fixed* cell protein which is indispensable for cell life or activity; (ii) *dispensable* reserve protein which can be called upon for energy and other purposes in starvation; (iii) *labile* reserve protein which can be readily turned out into the blood stream to maintain the plasma protein concentration. In hæmorrhage or in protein starvation such outflow from the tissues into the plasma takes place. If plasma proteins are given intravenously they can supply all the tissue needs for protein, and food protein can be temporarily dispensed with; this observation suggests that plasma protein can be readily incorporated into the tissues. Proteins taken by mouth after hydrolysis in the intestine are, of course, readily built up into plasma or tissue protein. One must conceive of rapid interchanges between the proteins in the liver, plasma, and the tissues generally.

Functions of Plasma Proteins.—(1) PROTEINS CONCERNED WITH BLOOD CLOTTING.—(i) *Fibrinogen.*—(*a*) The average fibrinogen content of plasma is 250 mg. per 100 c.c. (range 190–330 mg-%). When blood is shed, fibrinogen is converted into fibrin by the newly formed thrombin (p. 140). The fibrinogen level is not influenced by fasting or by the protein content of the diet, so it obviously has a "priority" claim on the tissue protein reserves.

(*b*) Plasma fibrinogen is raised in many acute *infectious diseases* [1]; this response is not specific; it occurs equally markedly in many other conditions, and is not due to the associated pyrexia. Thus pyrexia produced by altered environmental conditions does not affect fibrinogen concentration; intravenous injection of typhoid vaccine, on the other hand, produces pyrexia which is accompanied by a marked sustained increase in fibrinogen. Plasma fibrinogen may be raised in infectious diseases in the absence of fever or of leucocytosis. Failure of the plasma fibrinogen to rise is generally of bad prognostic significance and may indicate impaired liver function.[2]

(*c*) Plasma fibrinogen level is one of the important factors which determine the *sedimentation rate* of the red corpuscles (p. 172).

(ii) *Other Proteins Concerned in Blood Clotting.*—The rôle of prothrombin, thrombin and plasma thromboplastin is discussed on pp. 139 *et seq.*

(2) PROTEINS CONCERNED WITH MAINTAINING PLASMA COLLOID OSMOTIC PRESSURE.—See p. 15.

(3) MISCELLANEOUS PROTEINS.—The rôle of some of the other plasma proteins that have been isolated (p. 133) are considered in the following places: hypertensinogen (p. 349); alkaline phosphatase (p. 1002); anterior pituitary hormones (p. 930).

(4) CARRIAGE BY PLASMA PROTEINS OF OTHER PLASMA CONSTITUENTS.—(i) Albumin adsorbs a number of substances, *e.g.* injected phenol sulphonephthalein, Evans' blue dye (used in plasma volume determinations (p. 9)), sulphanilamide and *blood-thyroxine* (p. 976).

[1] Ham and Curtis, *Medicine*, 1938, *17*, 413.
[2] Plasma fibrinogen is also increased after trauma (p. 149).

(ii) An α-globulin absorbs *bilirubin*, the resulting compound being responsible for the indirect Van den Bergh reaction (p. 188). A β_1-globulin combines with *iron* (p. 208) and other metals. The globulins also combine with cholesterol, phospholipids, vitamin-*A*, and steroid hormones.

COAGULATION OF BLOOD.[1] VITAMIN-*K*. HÆMOSTASIS. HÆMORRHAGIC STATES [2]

Blood Coagulation.—The essential reaction in coagulation of the blood is the conversion of the soluble protein fibrinogen into the insoluble protein fibrin by means of an enzyme thrombin. Fibrinogen exists in the circulating blood as such; thrombin does not, but is formed from an inactive circulating precursor prothrombin when the blood is shed. The activation of prothrombin depends on the presence of Ca^{++} ions and of certain organic activators of obscure composition labelled thromboplastins which are derived from damaged tissues and disintegrating platelets and from the plasma itself. The formation of prothrombin (in the liver) depends on the absorption from the bowel of adequate amounts of vitamin-*K*. It should be remembered that clotting of blood is not the only reaction responsible for arrest of hæmorrhage (*hæmostasis*).[3] Hæmostasis also depends on certain reactions of the injured vessels.

It is essential that the blood should normally be kept fluid while circulating in the blood vessels, though clotting when shed. The fluidity of the blood in the body depends (*a*) on the special physical properties of the intact vascular endothelium, (*b*) on the flow of the blood (*i.e.* on its not stagnating in any part of the vascular bed), and (*c*) the presence in the blood of a natural anticoagulant agent, heparin. Clinical states of excessive bleeding (*hæmorrhagic states*) may be due to (i) impaired coagulability of the blood due to some abnormality in the complex physico-chemical system concerned in clotting, or (ii) alterations in the vessel walls preventing them contracting down after local injury. Abnormal coagulation of the blood within the blood vessels (*intravascular thrombosis*) is generally due to alterations in the vascular endothelium combined with slowing down of the rate of flow; it is rarely the result of a primary alteration in the blood clotting mechanism.

It is easy to give a simple account of the processes involved in blood clotting as outlined above, thus:

$$\begin{array}{l} \text{Prothrombin} + Ca^{++} + \text{Thromboplastins} \leftarrow \begin{cases} \text{Damaged Tissues} \\ \text{Blood Platelets} \\ \text{Plasma} \end{cases} \\ (\textit{inactive}) \\ \qquad\qquad \downarrow \\ \qquad\qquad \text{Thrombin} + \text{Fibrinogen} \\ \qquad (\textit{active enzyme}) \\ \qquad\qquad\qquad \downarrow \\ \qquad\qquad\qquad \text{Fibrin.} \end{array}$$

[1] Macfarlane, *J. clin. Path.*, 1948, *1*, 113.

[2] Quick, *Hæmorrhagic Diseases and Hæmostasis*, Springfield, 1942; *Physiology and Pathology of Hæmostasis*, 1951. Astrup, *Advances in Enzymology*, 1950, *10*, 1.

[3] Not to be confused with *homeostasis*, a word coined by Cannon to signify maintenance of the constancy of the internal environment.

Circulating heparin, which acts as an antithrombin, is neutralized by the newly formed thromboplastin.

This simple scheme, which is so easy to remember, accounts for most of the main facts of blood coagulation but does not account for all. Any attempt to account for all the facts reveals the coagulation process as it really is, namely as an elaborate series of physico-chemical reactions involving some of the proteins and certain other of the constituents of a complex colloidal system, namely the plasma.

Rôle of Calcium.—Ionic calcium is essential for clotting. Clotting is prevented (i) by addition of K oxalate which gives rise to a precipitate of insoluble Ca oxalate, or (ii) by addition of Na citrate which forms a soluble compound, Na calcio-citrate; but as the calcium is bound with the citrate radical it is no longer in the ionic form. For special purposes sodium fluoride is added to prevent clotting since it has the additional property of inhibiting blood enzymes. Decalcified blood or plasma will clot on addition of adequate amounts of a soluble ionizable calcium salt, *e.g.* $CaCl_2$. The normal serum Ca level (10 mg. total, 5 mg. ionized per 100 c.c.) is more than adequate for optimal speed of blood coagulation. In hypercalcæmia due to parathyroid excess (p. 1006) intravascular clotting may occur just before death. Ca^{++} ions are necessary for the activation of prothrombin but not for the action of thrombin on fibrinogen.[1]

Fibrinogen is a protein with a long thin thread-like molecule (dimensions, 900×33 Å.U.); its molecular weight is nearly 450,000; its isoelectric point is *p*H 5·5; and as is the rule with proteins, it is readily precipitated at this *p*H after reducing the electrolyte content of plasma by preliminary dialysis or dilution with distilled water. It is precipitated by heating to 47° C. It is probably formed in the liver; exclusion of the liver leads to a rapid fall in plasma fibrinogen concentration, indicating that it is rapidly utilized, though where or for what purpose is unknown. Fibrinogen is converted by thrombin into fibrin. Cases have been recorded in which the blood was incoagulable owing to the complete congenital absence of fibrinogen.[2]

Fibrin.—In clotting blood, fibrin is laid down as a network of fine threads which entangle the formed elements of the blood, principally the red cells. The freshly formed threads are extremely adhesive, sticking to each other, to the blood cells, to the tissues, and to certain foreign surfaces; this adhesiveness makes the clot an effective hæmostatic agent. Freshly shed blood sets in a soft jelly-like mass; gradually this clot contracts down (*retracts*) to about 40% of its original volume squeezing out *serum*[3]; plasma clots can contract down to as little as 10% of their original volume. The final clot is tougher and more solid and elastic and is presumably a more efficient bung to damaged vessels. Clot retraction is impaired if the platelets have been artificially removed or in disease conditions with a low platelet count, although the speed of clotting is not prolonged (p. 159); it is not known how the platelets promote clot retraction. Clots formed in the tissues

[1] *Oxalate plasma* is prepared as follows. Blood withdrawn from a blood vessel into a receptacle containing potassium oxalate (to a final concentration of 0·1%) is centrifuged and the supernatant plasma (*oxalate plasma*) is removed.

[2] Prunty, *Brit. J. exp. Path.*, 1946, *27*, 200.

[3] The essential difference between plasma and serum is that plasma contains fibrinogen and is coagulable while serum (having been squeezed out from the clot) contains no fibrinogen and is not coagulable.

have ultimately to be disposed of as healing takes place; the dissolution of the clot—*fibrinolysis*—is due to the action of proteolytic enzymes in the plasma which become activated.

Thrombin.—The properties of this substance will be considered here; its mode of formation in shed blood is discussed below (p. 142). The most purified specimens of thrombin are very potent and can coagulate at least 600 times their weight of fibrinogen. Thrombin is an albumin, with a molecular weight of 75,000; it contains no P or Ca; its properties are those of an enzyme. The velocity of the thrombin-fibrinogen reaction is accelerated by increased concentration of thrombin (Fig. 74). The reaction proceeds slowly at low temperature, accelerates with rising temperature to an optimum

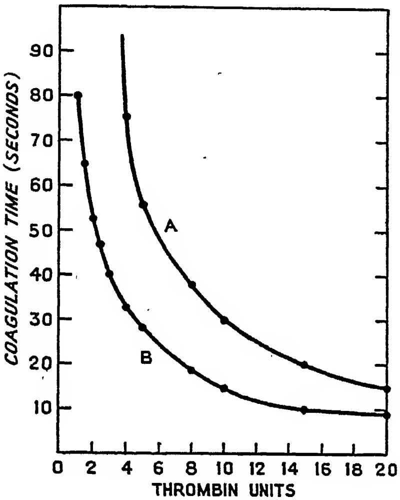


FIG. 74.—Relationship of Concentration of Thrombin to Clotting Time of Oxalated Plasma (A) and a Fibrinogen Solution (B). (McFarlane, *J. clin. Path.*, 1948, *1*, 126.)

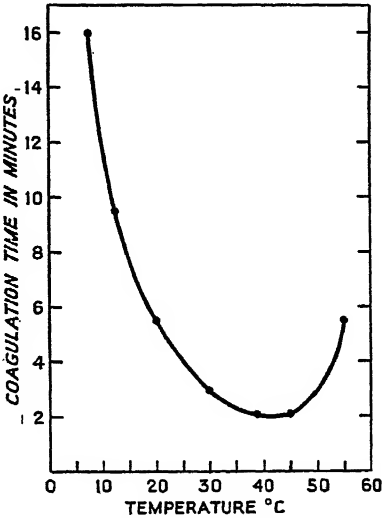


FIG. 75.—Effect of Temperature on Coagulation Time of Normal Whole Blood. (Macfarlane, *J. clin. Path.*, 1948, *1*, 126.)

at 40° C. and then declines; thrombin is destroyed by heating to 60° C. (Fig. 75).

Exactly what thrombin does to fibrinogen to make its molecule insoluble in water is unknown; presumably it alters its molecular configuration. Thrombin has been thought to be a proteolytic enzyme but this is uncertain. It must be emphasized that thrombin can act in the complete absence of Ca^{++} ions, *i.e.* it clots oxalated plasma; Ca^{++} ions thus play their part in the earlier stages of the coagulation process.[1]

Prothrombin.—This inactive precursor of thrombin is a globulin which is present in the circulating blood in a concentration of about 40 mg. per

[1] *Intravascular Injection of Thrombin.*—If large amounts of thrombin are injected intravenously (*e.g.* 3 mg. into a 2 kg. rabbit), death occurs within a few minutes from extensive intravascular clotting, especially in the great veins and heart. *No immunity* against thrombin can be produced by repeated injections of minimal amounts of thrombin.

100 c.c. of plasma. It is precipitated at pH 5·6 and destroyed by heating at 60° C.

Prothrombin is formed in the liver; its concentration in the plasma falls rapidly after hepatectomy. Its formation depends on an adequate supply and satisfactory absorption of *vitamin-K* (p. 151). In new-born babies the plasma prothrombin is considerably lower than in older children (p. 153).

Inception of Blood Clotting.—Thromboplastins.—The following observations are relevant to a consideration of this question:

(i) Blood carefully withdrawn from a blood vessel without contamination with tissue fluid does not clot if kept in a container with a surface which is *not water-wettable*, *e.g.* paraffin-lined or made of a plastic like *silicone.* The addition to such a blood of a watery extract of almost any tissue (especially brain, lung, testis, *platelets*) promotes rapid clotting. Thus circulating prothrombin in the presence of Ca^{++} ions remains inactive; it is activated by the complex materials which are present in tissue extracts and are called *thromboplastins* (*i.e.* " clot promoters ").

(ii) Blood collected as above (*i.e.* free from tissue fluid) but brought in contact with a water-wettable surface, rapidly clots. The greater the surface of contact, the greater is the rate of clotting. It is presumed that thromboplastins are liberated by contact with such a surface: (*a*) from the *platelets* which rapidly disintegrate in such circumstances, and (*b*) from an inactive precursor in the plasma (a view supported by observations on hæmophiliac blood (p. 150)).

(iii) When bleeding occurs following an injury, thromboplastins are liberated from *all* the sources enumerated above, namely damaged tissues, disintegrating platelets, and altered plasma.

(iv) Nature of Thromboplastins.—This question is unsettled; probably natural thromboplastin is an enzyme protein combined with a lipide (resembling *cephalin*) which *activates* the enzyme. *Aqueous* extracts of tissues contain the protein fraction; *ether* extracts the lipide fraction; the separated fractions are less effective than the entire complex. A very potent thromboplastin has been extracted from the venom of Russell's viper; added to plasma (even hæmophiliac plasma) it produces rapid clotting; but it does not clot plasma from which the lipide has been removed. The venom thus resembles the protein (enzyme) component of natural thromboplastin in needing the presence of an activator (lipide) for its effective action. Tissue extracts yield the whole active lipo-protein; platelets may yield the lipide factor. In shed *plasma* the protein-enzyme is probably liberated from an inert precursor and is then suitably activated by the natural lipides of the plasma.[1]

(v) Mode of Action of Thromboplastin.—In the presence of Ca^{++} and thromboplastin, prothrombin is activated to thrombin. As the thrombin molecule is smaller than that of prothombin and also contains no P or Ca, there is no question of the reaction being a simple additive one. The enzyme component of thromboplastin is believed to split off part of the prothrombin molecule thus converting it into active thrombin. [Other proteolytic

[1] It has been suggested that there is a plasma fraction (*thrombocytolysin*) which destroys (lyses) platelets thereby liberating their contained lipoid clotting factor. This factor is deficient in hæmophilia.

enzymes—*e.g.* crystalline trypsin—produce coagulation in the same way; and, as might be expected, agents which inhibit proteolysis inhibit blood coagulation.]

In addition thromboplastin *neutralizes the circulating heparin*—which is an antithrombin—permitting the newly formed thrombin to act.

Thrombin is probably formed in small quantities naturally in the circulating blood. It is immediately antagonized by heparin; in addition it combines with and is inactivated more gradually by an albumin fraction, consequently called antithrombin.[1]

Simple Coagulation Time.—This is most easily measured by finding the time taken for blood (collected by venepuncture or capillary puncture) to clot in fine capillary tubes; the absolute values are of little interest, but comparative values for normal blood and the blood under investigation are most useful. Prolongation of simple coagulation time is evidence of *gross* impairment of blood coagulability. Such studies show that most of the coagulation factors are present in considerable excess for the conditions used in the measurement; thus coagulation time is unaltered until (i) fibrinogen is almost completely absent, (ii) prothrombin is reduced to less than 10% of normal or (iii) total serum Ca is reduced below the level at which tetany occurs (5 mg. per 100 c.c.); in other words the *dominant factor normally controlling the clotting time of whole blood is the concentration of thromboplastin.*

Clinical Estimation of Prothrombin.—The principle employed is as follows: Oxalate plasma is prepared; more than enough ionized Ca salt and thromboplastin (*i.e.* brain or lung extract) are added to produce optimum conditions for the activation of the prothrombin to thrombin. It is assumed that the concentration of fibrinogen and of all other accessory factors is normal. The rate of clotting induced in these circumstances depends on the amount of thrombin formed and therefore on the *initial prothrombin concentration.* The time of clotting so produced is called the *accelerated clotting time* or the (Quick) *prothrombin time* (after Quick, whose technique is commonly used); a prolonged clotting time means a lowered prothrombin concentration. Curves have been constructed (Fig. 76) showing the relationship between prothrombin time and the prothrombin concentration (as a percentage of the normal); a prothrombin percentage of 100 represents the *normal* concentration in plasma. In Fig. 76 100% corresponds to a clotting time of 25 seconds. The actual normal clotting time obtained varies with each sample of brain extract and with some samples may be as low as 12 seconds. Pathological samples should therefore always be compared with normal controls. It should be noted that the prothrombin may fall to nearly

[1] *Anticoagulants.*—Coagulation may be prevented or delayed as follows:

(1) Prevent the blood from coming into contact with a water-wettable surface (*e.g.* collect in a paraffin-coated vessel), or rapidly cool blood to 0° C. These measures prevent the platelets and corpuscles from disintegrating, and so delay the formation of thromboplastic substances (p. 142).

(2) Remove free calcium ions from the solution (p. 140).

(3) Addition of neutral salts, *e.g.* half saturated Na_2SO_4. This dilutes the blood and apparently impedes the interaction of the various substances concerned in clotting.

(4) Use of *heparin* or related substances (*e.g.* hirudin); (injection of *peptone* leads to liberation of heparin) (p. 144).

(5) Use of *dicoumarol* (p. 154).

Clinically, to keep blood fluid for transfusion purposes, *sodium citrate* is generally employed. To prevent or retard blood clotting *in vivo* heparin or dicoumarol is used.

half with only a small prolongation of clotting time ; when the prothrombin concentration falls below 50% the clotting time is greatly prolonged.[1]

Heparin.[2]—Heparin was first isolated from the liver (hence its name) and shown to be a powerful anticoagulant substance ; it was subsequently demonstrated in extracts of many other organs, *e.g.* lungs. Heparin is probably normally *secreted* by a scattered widely distributed system of connective tissue cells called the *mast cells* ; it helps to maintain the normal fluidity of the blood within the vascular bed.

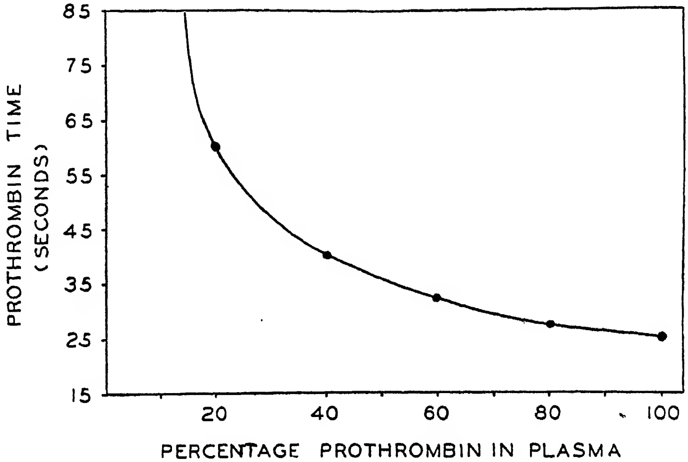


FIG. 76.—Relation between Prothrombin Time (time taken for oxalate plasma to clot on addition of Ca^{++} ions and tissue thromboplastin) and Prothrombin percentage (100=*normal concentration*) in Plasma. (Kark *et al., Quart. J. Med.*, 1940, *9*, 251.)

Heparin inhibits blood coagulation both *in vitro* and *in vivo* ; it acts slightly by preventing the activation of prothrombin to thrombin but mainly by neutralizing the action of thrombin itself. 1 mg. of a purified heparin preparation added *in vitro* to 500 c.c. of freshly withdrawn cat's blood kept at 0° C. may prevent its coagulation for 24 hours. Chemically heparin is

[1] The above account has been deliberately over-simplified for the sake of relative clarity. The following observations are worth mentioning in a footnote :

Prothrombin probably consists of two components, one labile and easily destroyed by heat and storage, the other relatively stable to heat and storage. The stable factor can be adsorbed on aluminium hydroxide and is the component which is decreased in the plasma of patients treated with dicoumarol. Plasmas treated in any one of the above ways all have long prothrombin times, but when heated or stored plasma (containing the stable factor) is mixed with a little normal plasma, or with adsorbed plasma, or with plasma from dicoumarolised patients (containing the labile factor) the prothrombin time is much shortened.

Owren described a case similar to hæmophilia, but with a long prothrombin time, which was restored to normal by the addition of a little normal plasma. He called the substance active in normal plasma, factor 5. Owren's factor 5 is probably the same as the labile component of prothrombin.

A *co-factor* is said to be necessary for the activating action of Ca^{++} ions.

[2] Jorpes, *Heparin*, London, 2nd edn., 1947. Quick, *Physiol. Rev.*, 1944, *24*, 297.

a *polysaccharide* built from glucosamine (amino-glucose) and glucuronic acid,[1] and containing many *sulphate* groups; it is thus related to mucoitin sulphate (of mucus) and chondroitin (of cartilage) (p. 839). A comparative study suggests that other high-molecular-weight polysaccharides containing several sulphate groupings possess heparin-like properties. Thus, if cellulose, glycogen, or starch are sulphated, they become highly anticoagulant (but, unfortunately, more toxic than heparin); monosaccharides or disaccharides, however, when similarly treated, remain inactive. Some of the artificial anticoagulants such as *Chicago blue* and similar dyes, *novirudin*, or the natural substance *hirudin*, extracted from the head of the leech, have a similar chemical constitution.

Certain dyes containing the grouping =NH, change colour when added to a heparin solution in a test tube; this property is called *metachromasia*; *toluidine blue*, for example, changes in colour to purple in the presence of heparin. [The reaction is specific for sulphuric acid esters ($R—OSO_3H$) and their salts if they have a high molecular weight.] Heparin is quantitatively precipitated from watery solution by toluidine blue and thus rendered inert; this reaction has been developed as a *quantitative* test for heparin. *Protamine* has a similar precipitating action to that of toluidine blue.

Mast Cells.—These cells, first described by Ehrlich, are widely distributed in many organs of all species from fish to man. They are found singly or in clumps; characteristically, they are arranged in close proximity to the walls of small blood vessels, and may even *replace the lining endothelium*. The cells contain numerous heparin granules which give a typical metachromatic purple reaction with toluidine blue.

(1) There is a correlation between the number of mast cells in a tissue, its SO_4'' content, and the amount of heparin that can be extracted from it. Thus the liver of sheep and oxen contains many mast cells and yields much heparin; the reverse is the case with the liver of the rat. On the other hand, the subcutaneous tissue in the rat is rich both in mast cells and heparin content. Considerable amounts of heparin are found in the walls of the large vessels, *e.g.* aorta and vena cava, where, too, the yield is related to mast cell concentration.

(2) In conditions in which the heparin content of the blood is raised (*e.g.* peptone shock, *v. infra*) the mast cells show loss of both granules and of metachromatic reaction.

(3) Metachromatic *extracellular* substance is also found under the intima and especially in the media of the aorta round the elastic fibres; it is also present in the substantia propria of the cornea. These findings, too, are related to a high heparin content.

Peptone Shock.—The intravenous injection of peptone produces a profound fall of blood pressure and incoagulability of the blood owing to an increase in heparin content; the mast cells appear exhausted from loss of granules (Fig. 77). The blood platelet count falls. The addition of toluidine blue to the blood reduces its coagulation time (by precipitating heparin). The heparin equivalent of the blood (judged by the toluidine blue reaction) in peptone shock is 3–6 mg. per 100 c.c. compared with a normal concentration

[1] $\left.\begin{array}{l}\text{Glucose}=CH_2OH(CHOH)_4CHO\\ \text{Glucuronic acid}=CH_2OH(CHOH)_4COOH.\\ \text{Glucosamine}=CH_2OH.(CHOH)_3CHNH_2.CHO.\end{array}\right\}$ For open chain structure, see p. 844.

of about 0·15 mg. Injection of peptone acts similarly in the liverless animal. It seems probable therefore that it acts on the mast cells generally, causing them to discharge their heparin store.

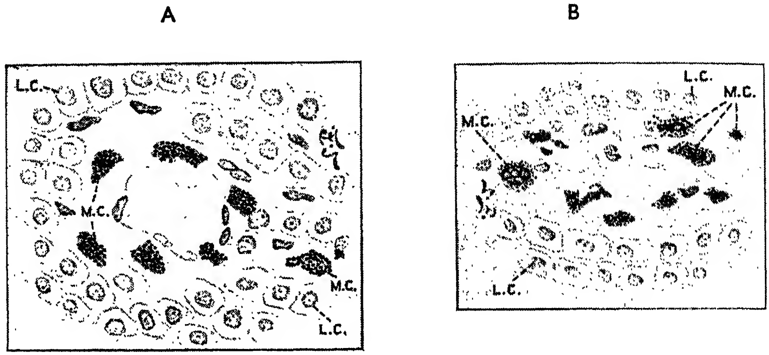


FIG. 77.—Liberation of Heparin from Mast Cells in Liver. (Wilander, *Skand. Arch. Physiol.*, 1928, *81*, *Suppl.* 15.)

A. Normal liver (dog) × 1150. Fixed with 4% basic Pb acetate and stained with 1% aq. toluidine blue. The mast cells (M.C.) are arranged round a small blood vessel and filled with well-marked granules of heparin. L.C. = hepatic cells, M.C. = mast cells.

B. Liver after peptone shock. The granules in the mast cells (M.C.) are diminished in number, smaller, and more faintly stained.

NORMAL RÔLE OF HEPARIN.—The facts presented make it likely that the mast cells normally secrete heparin into the blood to help maintain its fluidity, partly by preventing the activation of prothrombin, but mainly by

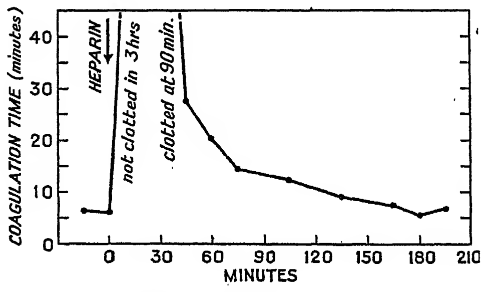


FIG. 78.—Action of Single Injection of Heparin on Blood Clotting Time. Normal subject (weight = 73 kg.).

At the arrow: inject 5000 units of heparin intravenously [1 unit = 0·009 mg. of pure heparin sodium salt; 5000 units = 45 mg.].
Normal clotting time = 5–10 minutes.
A blood sample withdrawn 15 minutes after the injection did not clot in 3 hours; a sample withdrawn after 30 minutes clotted in 90 minutes. The clotting time returned to normal in about 3 hours. (Argent, Gilliatt, and Slack, *Middlesex Hosp. J.*, 1948, *48*.)

neutralizing any thrombin that might be formed. When blood is shed, heparin is put out of action by the thromboplastins which are liberated from damaged tissues, so permitting the rapid conversion of prothrombin into thrombin and the subsequent action of thrombin on fibrinogen.

Use of Heparin.—Heparin can be employed to keep blood fluid *in vitro* for purposes of analysis or *in vivo* in various experimental conditions. It may be added to the donor blood in blood transfusions.

Heparin may be injected intravenously in cases of commencing or developed thrombosis, *e.g.* in the deep veins of the leg after operations, or in the coronary or cerebral vessels, in the hope of preventing the spread of the intravascular thrombosis; favourable clinical results have been reported. Fig. 78 shows that a single intravenous injection of 5000 units (=45 mg.) of heparin produces a striking effect on the blood clotting time; samples collected after 15 minutes may have a clotting time which exceeds 3 hours (normal, 5–10 minutes). As heparin *rapidly* disappears from the blood the effects pass away in a few hours. In the case illustrated by Fig. 79 repeated doses of heparin produced a sustained and progressive *increase* in the clotting

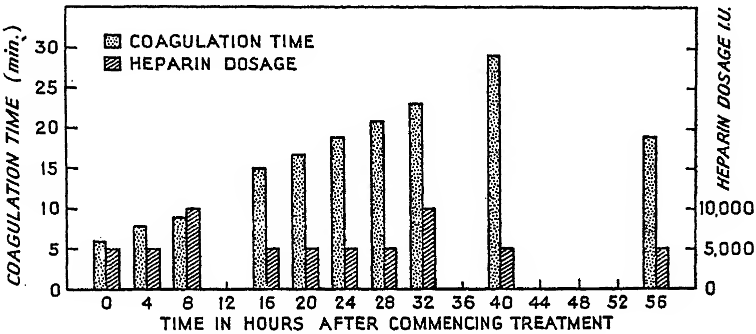


Fig. 79.—Action of Repeated Intravenous Injections of Heparin on Blood Clotting Time in Case of Venous Thrombosis.

The patient, aged 65, weight 51 kg., had a carcinoma of the rectum removed by the perineal route. On the eighteenth day he complained of pain in the right calf; there was pitting œdema of the right ankle and deep calf tenderness, indicating thrombosis in the deep veins of the calf. Heparin was given intravenously, generally at four-hourly intervals (=30,000 units in all) in the first 24 hours. During the second 24 hours the total dose was 20,000 units. The blood clotting time (normal=5–10 minutes) was progressively raised to 25 minutes, *i.e.* the blood was *much less coaguable.* On the third day only 5000 units were injected. Recovery was uneventful. (Argent, Gilliatt, and Slack, *Middlesex Hosp. J.*, 1948, *48.*)

time. There are, however, marked individual variations in the response observed; probably heparin is best given clinically by means of a continuous intravenous drip transfusion, the amount administered being adjusted to maintain the desired level of blood clotting time. Heparin therapy has the advantage over dicoumarol (p. 154) that its effects wear off quickly when administration is stopped and so a dangerous degree of incoagulability of the blood can be avoided or soon remedied.

Intravascular Thrombosis.[1]—The pathologists use the word *thrombus* to describe a "peculiar clot formed during life in the streaming blood"; the responsible process called *thrombosis* is distinguished from the usual extravascular clotting, or clotting in wounds, or clotting occurring in the blood vessels after death. Fundamentally "thrombosis", like "clotting", consists of fibrin formation; but there are differences described below which are due to the fact that the blood is streaming and not stationary.

[1] Marple and Wright, *Thromboembolic Conditions and their treatment with Anticoagulants* 1950, Springfield, Illinois.

Thrombosis always begins by deposition on the wall of a blood channel of *masses of platelets* which grow by adhesion of other platelets as they flow by; the laminæ of platelets (which fuse together and lose their identity) stand out as layers running transversely to the blood stream; passing leucocytes adhere to their borders (" like flies on sheets of sticky flypaper ") (Fig. 80). The platelets liberate thromboplastins " so that filaments of fibrin spread out from them on all sides and meeting with filaments from the next

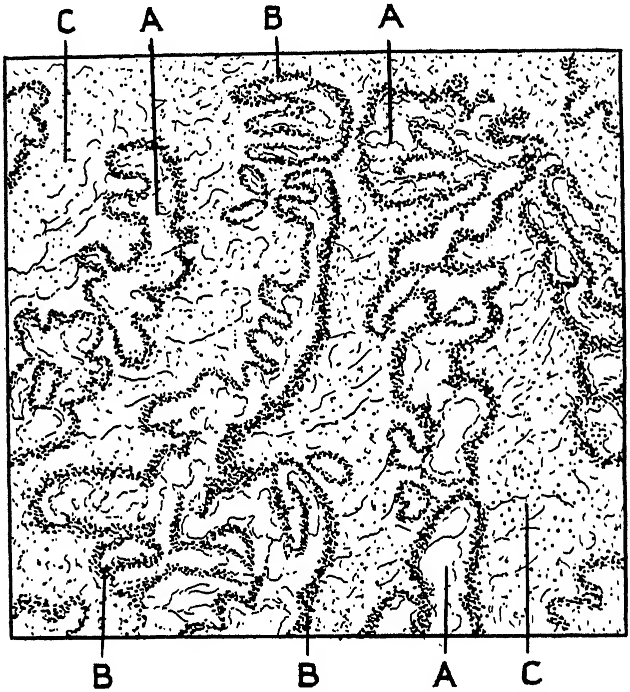


FIG. 80.—Structure of Intravascular Thrombus.

The pale translucent laminæ are the fused masses of platelets (A). The dense layer of deeply staining cells adherent to the laminæ are leucocytes (B). Between the platelet laminæ is a fibrin network entangling disintegrating red cells and some leucocytes (C). (McCallum, *Text Book of Pathology*, W. B. Saunders, Philadelphia and London, 1941).

lamella hang in festoons between them." The lamellæ of platelets thus " braced together by fibrin " entangle masses of red cells so that finally a solid mass of peculiarly constructed " clot "—in fact, the " thrombus "—is formed. (McCallum.) The red cells disintegrate and lose their hæmoglobin; the initial red thrombus as it ages becomes yellowish grey; but newly formed thrombi added to it will be red.

Thrombi form most readily where there is (i) local damage to the vascular endothelium, and (ii) slowing down of the blood stream, *e.g.* in small leg veins (*infra*), on atheromatous patches in small arteries, on damaged valves in the heart, or in the auricular appendage in auricular fibrillation. It is

surprising however to find thrombi forming on the damaged wall of the aorta where "it might seem that the pulsating torrent of blood would allow no chance for the deposition of pioneer platelets."

POST-OPERATIVE THROMBOSIS, THROMBOPHLEBITIS OR DECUBITUS THROMBOSIS.—After surgical operations (especially those involving the abdomen) or childbirth, and in patients confined to bed for long periods, thrombosis may occur in the leg veins; the condition is called post-operative thrombosis, decubitus thrombosis (because the recumbent position is an important causal factor), or thrombophlebitis (because the thrombosis is accompanied by changes in the vein wall). The thrombosis sets in within 3 weeks of the operation; it begins in the veins of the calf muscles and the plantar region and spreads upwards, in some instances ascending to the popliteal, femoral, or even iliac veins.[1] In a small proportion of cases of thrombosis fragments of clot become detached (*embolus*) and settle in and obstruct the pulmonary vessels producing infarction of the lungs; sometimes the cerebral vessels and rarely the coronary vessels become blocked in this way.

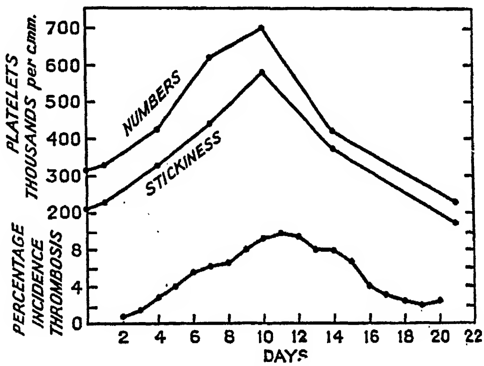


FIG. 81.—Relationship between Changes in Blood Platelet Count and "Stickiness" and the incidence of Post-operative Thrombosis. (After Helen Payling Wright.)
"Stickiness" of platelets in arbitrary uuits.

Mechanism of Decubitus Thrombosis.—This condition is due partly to changes in the circulation and in the properties of the blood and partly to local injuries to the leg veins.

(i) The circulation in the veins of the legs and trunk (but not in the arms) is considerably slowed down after operations. The normal venous return depends on muscular contraction and respiratory movements; after abdominal operations the legs are moved very little, and the movements of the diaphragm may be hampered by a tight abdominal bandage or by flatus or inhibited by the pain of the abdominal incision. Thrombosis is much rarer after operations on the upper part of the body.

(ii) There are changes in the composition of the blood owing to the general tissue response to the trauma: (*a*) The plasma fibrinogen concentration is raised; this may increase the sedimentation rate of the red cells (p. 173). (*b*) The platelet count is raised; the platelets also become more "sticky" and so more liable to adhere to the lining of the blood vessels. There is a direct relationship between the extent of these blood changes and the incidence of intravascular thrombosis (Fig. 81).

(iii) Sepsis may be a factor; thus most strains of *Staphylococcus aureus* produce a toxin which rapidly clots human blood.

(iv) The calf veins may be damaged owing to the limbs lying

[1]See legend to Fig. 79 for a clinical report of a typical case.

limply on the operating table or in bed; thromboplastins are liberated locally promoting thrombin formation.

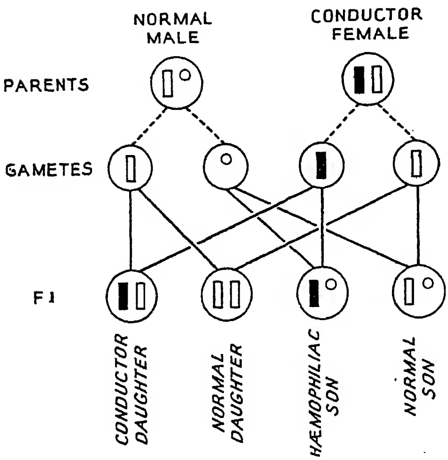


FIG. 82.—Transmission of Hæmophilia by Conductor Female to Hæmophiliac Son and Conductor Daughter.

▯ =normal x chromosome.

▮ =x chromosome responsible for hæmophilia.

○ =y chromosome.

The sex chromosomes in the *female* are similar in appearance and are labelled x and x; in the male the sex chromosomes are x, combined with a minute and negligible chromosome labelled y. The gene responsible for hæmophilia is present in the x chromosome; in the presence of another normal x chromosome the gene acts as a *recessive, i.e.* the individual has no signs of hæmophilia (but can transmit the disease); certain constituents of the normal x chromosome may be responsible for this. When the ovum of a conductor female undergoes its reduction division, the two resulting cells differ: one contains a normal x chromosome ▯, the other a hæmophiliac x chromosome ▮. At fertilization, if x of the ovum unites with x of a sperm, the offspring is a female if x of the ovum unites with y of a sperm the result is a male. If a hæmophiliac x ▮ (of a female) unites with a y (of a male) the hæmophiliac gene on x ▮ is *unantagonized* by a normal x ▯ and the resultant is a *hæmophiliac son.* If a hæmophiliac x ▮ unites with a normal x the result is a conductor daughter (who can transmit the disease but has no signs of it). Offspring not containing hæmophiliac x ▮ are normal and their progeny are normal.

The mode of formation and spread of the thrombus was described above (p. 148).

The use of heparin and dicoumarol to prevent the onset or spread of thrombosis is considered on pp. 147, 154.

Hæmophilia[1] is an inherited anomaly usually transmitted by females, who themselves show no symptoms, to males who manifest signs of the disease (Fig. 82). The condition is characterized by a *marked increase in the coagulation time.* Blood should be collected for examination by venepuncture; normal blood under these conditions (*i.e.* in the absence of thromboplastins) clots in 5–10 minutes, while hæmophiliac blood may take from 1–12 hours. Severe bleeding occurs after injuries of any kind, even of the most trivial character. There is no tendency, however, to *spontaneous* hæmorrhage, thus distinguishing the condition from purpura (p. 157).

The abnormality responsible for hæmophilia is probably a deficiency of a plasma "thromboplastin." The following points should be noted:

(i) The concentrations of fibrinogen, Ca^{++} ions, prothrombin, platelets, and heparin are normal.

(ii) The patient's prothrombin is activated normally by adding excess Ca^{++} ions and tissue thromboplastins.

(iii) Intravenous injection of 100 c.c. of *normal* blood (but not of hæmophiliac blood) shortens the coagulation time to about normal.

(iv) Injection of Cohn's protein Fraction 1 (p. 136) (containing fibrinogen

[1] Howell, *Bull. N.Y. Acad. Med.*, 1939, *15*, 3. Minot *et al.*, *J. clin. Investig.*, 1945, *24*, 704; 1946, *25*, 870, 876. Pavlovsky, *Blood*, 1947, *2*, 185. Craddock and Lawrence, *ibid.*, 505.

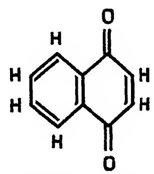
and globulins) extracted from *normal* plasma but *not from hæmophiliac plasma* has a similar effect. As the fibrinogen concentration of hæmophiliac blood is normal, a globulin component of the added plasma or of Fraction 1 is presumably the active agent; likewise hæmophiliac plasma clots normally *in vitro* on adding normal plasma free of platelets, fibrinogen, or prothrombin.

It seems that in hæmophilia a normal plasma thromboplastin is missing and that it is essential for satisfactory thrombin formation. Presumably this plasma thromboplastin is associated with the globulin fraction. As a helpful measure to cope with local and *accessible* bleeding in hæmophilia a preparation rich in thromboplastins (like the venom of the Russell viper) may be applied on a dressing at moderate pressure and usually produces satisfactory clotting.

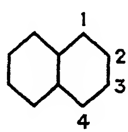
Vitamin-K and Blood Clotting.[1]—Vitamin-*K* is a complex *naphthoquinone* derivative which is indispensable for the synthesis of *prothrombin* by the liver; in the absence of the vitamin the plasma prothrombin concentration is reduced, the coagulation time of the blood is correspondingly prolonged and serious hæmorrhages may occur.

Distribution and Chemistry.—Vitamin-*K* is widely distributed in nature; thus it is found in *traces* in green vegetables, cereals, and animal tissues generally. It can be synthesized by many *bacteria* including those normally present in the human intestine (*e.g. B. coli*); the bacterial flora probably provide an adequate supply of vitamin-*K* in man. Vitamin-*K* deficiency from dietary abnormalities is very rare in man.

The formula of 1 : 4 naphthoquinone (which contains an O atom at the 1, and 4, positions) and the method of numbering the positions in the ring are indicated below:



1 : 4 napthoquinone



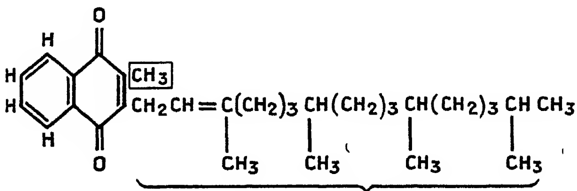
Natural vitamin-*K* is 2-methyl-3-phytyl-1 : 4 naphthoquinone (Fig. 83). It is *soluble in fat-solvents* but *insoluble in water*; it can therefore only be administered by mouth or intramuscularly. Many substances of related chemical constitution with vitamin-*K* activity have been synthesized. The most potent (with three times the activity of natural vitamin-*K*) is 2-methyl-1 : 4 naphthoquinone; it is interesting to note that a closely similar substance chemically (2-methyl-3-*hydroxy*-1 : 4 naphthoquinone (*phthiocol*))is almost inert biologically. *Water-soluble* analogues have also been prepared (*e.g.* 4-amino-2-methyl-naphthol-hydrochloride) which can be administered intravenously.

Absorption in Intestine.—Absorption of vitamin-*K* from the small intestine only occurs in the presence of adequate amounts of *bile salts* (which are also necessary for the absorption of other fat-soluble substances like vitamin-*D* and the fats of the food (p. 798)).

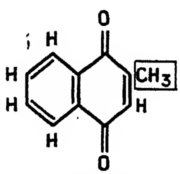
[1] Butt and Snell, *Vitamin*-K, Philadelphia, 1941. Almquist, *Physiol. Rev.*, 1941, *21*, 194.

The vitamin is absorbed into the lacteals and passes into the thoracic duct[1]; on reaching the liver it participates in the processes leading to the synthesis of prothrombin; the vitamin is *not* incorporated in the prothrombin molecule which is a protein containing no naphthoquinone. If bile is excluded from the bowel, vitamin-*K* absorption does not occur; if the liver is damaged or extirpated, prothrombin formation is decreased or stops altogether.

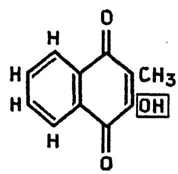
Clinical Conditions associated with Vitamin-K Deficiency.—(1) OBSTRUCTIVE JAUNDICE.—In this condition bile is excluded from the



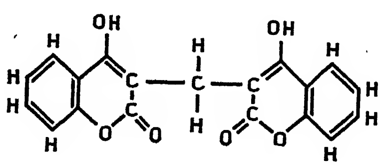
Natural Vitamin-*K*.



2-methyl-1 : 4 naphthoquinone.



2-methyl-3-hydroxy-1 : 4 naphthoquinone (phthiocol).



Dicoumarol.
(3, 3′—methylene—bis—(4 hydroxy)—coumarol.)

FIG. 83.—Chemical Structure of Vitamin-*K* and related Substances.

bowel owing to obstruction somewhere in the biliary passages. The absence of bile salts from the bowel prevents vitamin-*K* absorption with the expected results of lowered plasma prothrombin,[2] prolonged clotting time, and hæmorrhages. In untreated cases operative procedures may involve death from uncontrollable bleeding. The state of the blood may be restored to normal (i) by giving bile salts alone or bile salts plus vitamin-*K* by mouth (Fig. 84); (ii) more rapidly by injecting vitamin-*K* intramuscularly; (iii)

[1] Mann and Higgins, *Blood*, 1950, *5*, 177.
[2] For estimation of plasma prothrombin clinically see p. 143.

improvement can be produced within 2 hours by *intravenous* injection of a water-soluble vitamin-K analogue.

(2) The syndrome of vitamin-K deficiency has been reported in cases of ulcerative *colitis*. The cause may be (i) failure of absorption as a result of the altered state of the intestinal wall and the associated diarrhœa, or (ii) an abnormal state of the intestinal bacteria leading to defective synthesis of

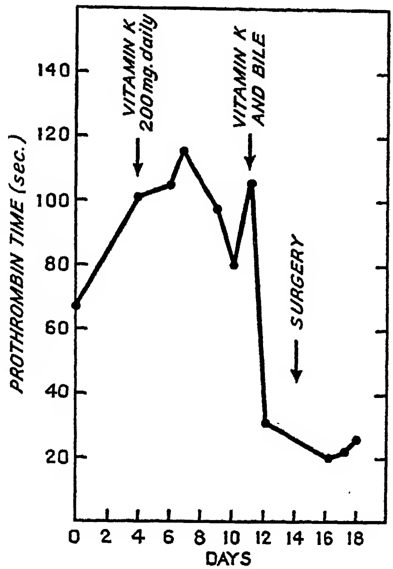


FIG. 84.—Action of Vitamin-K by Mouth with and without Bile in Case of complete Obstructive Jaundice.

Vertical axis: Prothrombin time (=accelerated clotting time) in sec. (*i.e.* clotting time of oxalated plasma after adding Ca^{++} and thromboplastin, Quick's method). Normal=25 sec. In this case of obstructive jaundice the prothrombin time was 70–100 sec. (*i.e.* the prothrombin concentration was less than 20% of normal (cf. Fig. 76)). Vitamin-K, 200 mg. daily by mouth was ineffective; but vitamin-K *plus bile* by mouth produced prompt restoration of plasma prothrombin to normal. (Butt and Snell, *Vitamin*-K, Philadelphia, 1941.)

the vitamin.[1] The syndrome may also occur associated with defective fat absorption in *sprue*.

(3) LIVER DISEASE.—In acute and subacute hepatitis and frequently in cirrhosis or malignant disease of the liver there may be failure of prothrombin formation with the usual sequelæ, but the condition does not respond to vitamin-K administration by any route (Fig. 85). This lack of prothrombin response has been used as a test of liver efficiency (p. 831).

(4) HÆMORRHAGIC STATES IN INFANTS.[2]—As already mentioned (p. 142) new-born babies commonly have plasma prothrombin concentrations which

[1] Kark and Souter, *Lancet*, 1940, ii, 1150.

[2] The infants also suffer from severe jaundice, marked anæmia, and widespread œdema. The relation of the vitamin lack to these other manifestations is unknown

are as low as one-third to one-sixth of normal; the deficiency is even more marked in premature babies. Usually the plasma prothrombin returns to normal during the second week after birth. A grave hæmorrhagic state has been described in infants in whom for some unexplained reason the plasma prothrombin falls to less than 1% of normal. If vitamin-*K* is given in large doses complete recovery of the blood may occur in less than 48 hours.[1]

DICOUMAROL.—This substance is related chemically to the naphthoquinone derivatives with vitamin-*K* activity as is shown by a study of Fig. 83. Because of this chemical resemblance, dicoumarol acts as an *anti*-Vitamin-*K* by the

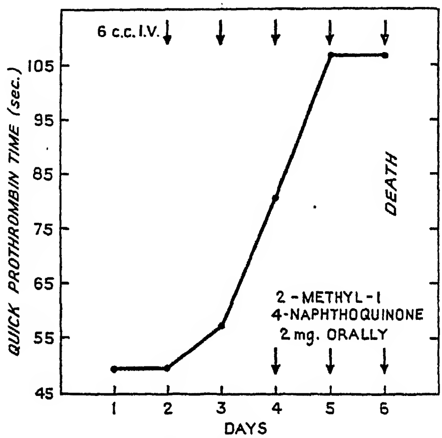


FIG. 85.—Absence of Prothrombin Response to treatment with Vitamin-*K* analogues in Acute Hepatic Insufficiency.

Case of acute yellow atrophy of the liver (acute hepatitis). Prothrombin time, 50 sec. (normal, 25 sec.) corresponding to prothrombin concentration of 30% of normal. Treatment consisted of:

(i) Intravenous injection of 6 c.c. daily of a water-soluble vitamin-*K* analogue (as shown by the upper arrows) for 5 days.

(ii) 2 mg. of 2-methyl-1 : 4-naphthoquinone daily by mouth for the last 3 days (lower arrows).

The condition steadily got worse. Prothrombin time rose to 105 sec. (=prothrombin concentration of under 10% of normal), before death occurred. (Kark and Souter, *Lancet*, 1940, ii, 1150.)

method of *substrate competition* (cf. p. 509). It is thought that in the liver dicoumarol replaces vitamin-*K* at the latter's normal site of action and thus prevents the vitamin from carrying out its normal physiological function. As dicoumarol cannot be utilized by the liver in the synthesis of prothrombin, the syndrome of vitamin-*K* deficiency develops, *i.e.* the plasma prothrombin level falls and blood coagulability is depressed. In this way dicoumarol acts as an anticoagulant, but only *in vivo*. It is active by mouth and one dose may produce an effect lasting for a week or more (Fig. 86). Dicoumarol has been used clinically, *e.g.* in phlebitis and coronary thrombosis. As the action of dicoumarol is so *prolonged*, repeated doses may produce a *cumulative* effect: the plasma prothrombin may fall so much that dangerous hæmorrhages may occur.[2] The drug should, therefore, only be used where there

[1] Dam, Tage-Hansen, and Plum, *Lancet*, 1939, ii, 1157.
[2] Thorsén, *Lancet*, 1947, i, 420.

are adequate facilities available for accurate frequent determinations of the plasma prothrombin level and for blood transfusion.[1]

Blood Platelets.—The red marrow contains giant cells (*megakaryocytes*), which may attain a diameter of 40 μ and contain an irregular ring of lobed nuclei. These giant cells throw out pseudopodia which pass through the wall of the sinusoids and constrict off part of their substance to form the blood platelets (*thrombocytes*). With Leishman's stain, the platelets appear as round or oval bodies 2·5 μ in diameter, with faint blue cytoplasm and distinct reddish-purple granules. If the stain is precipitated on the granules an appearance simulating a nucleus is produced. The normal platelet count is 250,000–500,000 per c.mm.[2]; in disease conditions the count varies as a

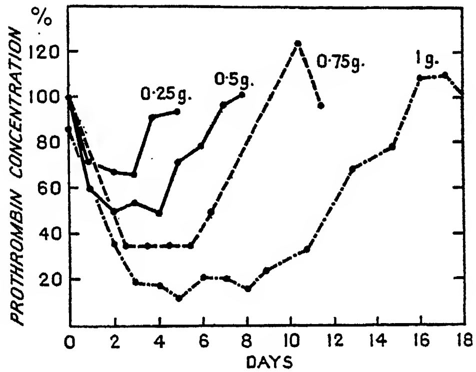


Fig. 86.—Action of Dicoumarol on Plasma Prothrombin Concentration.

Vertical axis: plasma prothrombin concentration. 100% = normal concentration; lower values = lower concentrations and therefore prolonged clotting time. The effects of taking by mouth 0·25, 0·5, 0·75 and 1·0 g. of dicoumarol are shown. (Redrawn from Lehmann, *Lancet*, 1942, i, 318.)

rule with the number of giant cells in the red marrow.[3] They are thus increased in myeloid leukæmia or in Hodgkin's disease (where they may exceed one million per c.mm.), and are diminished after large doses of X-rays, which destroy the marrow, or in severe infections of the marrow. The *duration of life* of the platelets is a few days.

The platelets are destroyed in the *spleen*; after splenectomy their number

[1] Burt *et al.*, *Brit. med. J.*, 1949, ii, 1250, have reported the results of clinical tests with bis-3, 3- (4-oxy coumarinyl) ether acetate (B.O.E.A.; trade name "tromexan"), a new coumarol substance which has the advantages of acting more quickly and being excreted more quickly than dicoumarol. Its use, however, requires the same rigid laboratory control.

[2] A platelet count is carried out as follows: The greased skin is punctured through a drop of sodium citrate solution which decalcifies the blood and preserves the platelets. The blood is transferred to a small waxed test-tube containing citrate solution with which it is thoroughly mixed. A drop of the greatly diluted blood is examined under a cover-slip on a glass slide, and the ratio of platelets to red cells is determined by counting 100–200 platelets and the red cells in the corresponding fields. The red cell count is also determined in the usual way. If the red count is 5 million per c.mm., and the red cell-platelet ratio is 10/1, then the platelet count is 500,000 per c.mm.

[3] In thrombocytopenic purpura, there may be few megakaryocytes in the marrow corresponding to the low platelet count. Sometimes the number of megakaryocytes may be normal; but the cells are then mostly immature and fail to form platelets.

increases, although only temporarily. In conditions of splenic over-activity (*hypersplenism*, p. 229) the platelets may almost disappear from the circulation. The platelet count is increased after all forms of *trauma* (p. 149).

FUNCTIONS.—(1) When blood is shed the platelets disintegrate and liberate a thromboplastin which is concerned with the activation of prothrombin to thrombin (p. 142). If a film of clotting blood is examined, it is noted that filaments of fibrin radiate from the vicinity of altered platelets. But the platelets are in no way indispensable for blood clotting; thus the number of the platelets is unaltered in hæmophilia, but the coagulation time is greatly prolonged (p. 150); in purpura, on the other hand, the platelets may almost disappear from the blood, but the coagulation time is within normal limits (p. 159). This latter observation indicates that excess thromboplastin is normally available in wounds from damaged tissue and the plasma

A B

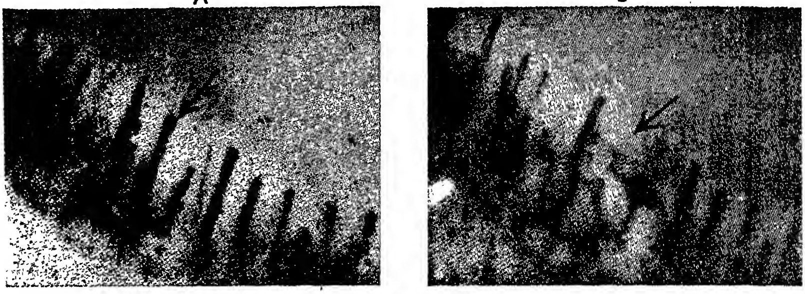


FIG. 87.—Response of Human Capillaries to Needle Puncture.

A. Normal capillary loops in the nail-bed of the human finger. The arrow indicates the approximate path of the needle.
B. The punctured capillary has disappeared as a result of contraction. (Macfarlane, *Quart. J. Med.*, 1941, *10.*)

itself. The platelet count is however related to the degree of *clot retraction* (p. 159).

The rôle of the platelets in the initiation of *intravascular thrombosis* is fully considered on p. 148.

(2) Diminution in the number of the platelets (*thrombopenia*, *thrombocytopenia*) is often associated with *purpura*; the significance of this observation is discussed on p. 159.

Hæmostasis.—As pointed out on p. 139 arrest of bleeding (hæmostasis) does not depend exclusively on coagulation of the blood; conditions of severe, apparently "spontaneous," hæmorrhage may occur clinically although the coagulability of the blood is normal. Hæmostasis depends in part on appropriate *vascular responses* at the site of injury. If the skin is *punctured* (*e.g.* with a needle) a number of capillary loops are damaged; "H" substance is liberated locally (p. 325), dilating the affected vessels, and bleeding takes place. After a few seconds the damaged vessels constrict (as seen when they are examined microscopically) and disappear from view (Fig. 87); bleeding ceases principally because the vessels have contracted. Under the conditions of these experiments the arrest of bleeding preceeds blood clotting, for, as has been pointed out, "accelerated blood clotting" *in vitro* (in the presence of

excess calcium ions and thromboplastins) takes 12–25 seconds; without such additions shed capillary blood clots in 1–2 minutes. The damaged capillaries remain shut for 20 minutes to 2 hours; during this period the blood that escaped initially has ample time to clot and the clot can become tough and firmly secured to the wound. Direct experiment in man shows that contracted capillaries may resist distension from internal pressures up to about 100 mm. Hg (p. 322); they can thus withstand the pressure in the neighbouring arterioles and small arteries. When the capillary circulation is finally re-established resumption of bleeding is prevented by the formed clot. If the edges of the wound have in the meantime come together owing to drying of the exudate further security against hæmorrhage is obtained.

Hæmorrhagic States.—The preliminary discussion (above) suggests that clinical hæmorrhagic states may be due to (i) defective blood clotting, (ii) defective capillary contractility, or (iii) the combined defects.

(i) In cases of *defective blood coagulation* a firm clot is not formed following an injury during the period of capillary contraction. When the capillaries finally open up once more, oozing will recommence, and can only be controlled by measures, either general or local, that restore blood coagulability. Clinically, failure of coagulation may be due (*a*) (very rarely) to lack of *fibrinogen* (p. 140); (*b*) to decrease of *prothrombin* (p. 151); (*c*) in *hæmophilia*, owing to lack of plasma thromboplastin (p. 150); (*d*) to lack of the lesser-known factors concerned in thrombin formation (p. 144). In these clinical states hæmorrhages may occur anywhere in the body "spontaneously"; it must be supposed that the capillaries in many regions are constantly being exposed to trivial and unnoticed trauma and that the resulting tiny blood-leaks are normally effectively sealed off. But when there is decreased blood coagulability from any cause the normal harmless leaks become noticeable or even dangerous hæmorrhages.

In hæmorrhagic states coagulating agents applied *sufficiently firmly* on the bleeding area may be useful temporarily; *e.g.* thromboplastins in hæmophilia (p. 151) or active preparations of thrombin. The dressing may consist of fibrin sheets or foam which do not need removal and are absorbed during healing.

(ii) *Defective capillary contractility.*—The clinical condition in which the capillary abnormality results in bleeding is known as purpura, which must now be considered.

Purpura.[1]—This is a condition with *normal blood coagulability* in which there is a tendency to "spontaneous" hæmorrhages, usually beneath the skin, from the various mucous membranes, and in internal organs. Purpura may be *symptomatic* or *primary*. Symptomatic purpura may result from various infections (*e.g.* infective endocarditis, typhus), from very many drugs in susceptible subjects (*e.g.* iodine, bismuth, ergot, quinine, sedormid), and in cachectic states (*e.g.* cancer). Primary (idiopathic) purpura occurs most often in children, has no constant associations with other maladies, and is occasionally congenital or hereditary. Severe purpura with hæmorrhages in the skin and from mucous membranes is called *P. hæmorrhagica*.

In purpura the blood coagulation time is, as stated, normal; but there is clear evidence of an *abnormal state of the capillary wall.* The following tests are used.

[1] Whitby and Britton, *Diseases of the Blood*, 6th edn., 1950, p. 336.

(1) Capillary Resistance.—If firm pressure (*e.g.* by inflating a blood-pressure cuff at 60 mm. Hg for 2 minutes) or a suction force (*e.g.* by the negative pressure used in "cupping") is applied to the skin, the local (and sometimes the distal) capillaries leak blood leading to the appearance of a crop of minute hæmorrhages (*petechiæ*); this abnormal response is evidence of diminished resistance (increased "fragility") of the capillary endothelium.

(2) Bleeding Time.—The ear is pricked and the escaping blood is washed away with saline or dried every 10 or 15 seconds with blotting paper;

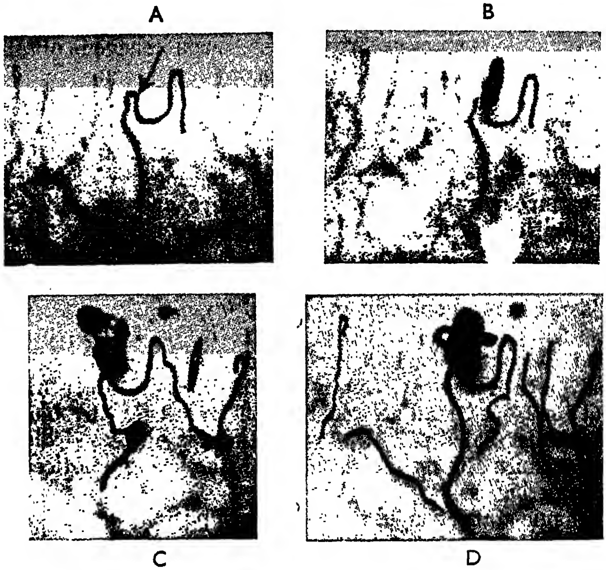


Fig. 88.—Distorted Capillaries and their Response in Primary (Athrombocytopenic) Purpura to Needle Puncture.

A. Note the abnormally formed capillary which is being punctured (needle track shown by arrow).
B. Immediately after puncture; hæmorrhage is taking place from the arteriolar limb of the capillary; the capillary has not contracted.
C. Two minutes after puncture.
D. Five minutes after puncture. During C and D blood had been escaping continuously from the needle track. Note that no capillary contraction has occurred. (After Macfarlane, *Quart. J. Med.*, 1941, *10*.)

normally bleeding ceases after 2–6 minutes. Under these conditions no blood clot can form locally and arrest of bleeding depends exclusively on capillary contraction. In purpura the damaged capillaries fail to close and thus bleeding may continue for 10–20 minutes or even longer.

(3) Skin Microscopy.—Examination in cases of *primary* purpura reveals that the skin capillaries are very irregular and distorted in form, sometimes branching; after puncture these vessels remain patent with the result that free bleeding proceeds from the needle track for several minutes (Fig. 88). In *symptomatic* purpura the capillaries are anatomically normal but (because of the presence of a toxic agent or other cause) they do not contract effectively in response to injury.

The facts just presented indicate the importance of the capillary defect in producing the hæmorrhages of purpura. Probably the capillary changes are localized, thus accounting for bleeding from restricted parts (*e.g.* skin and mucous membranes).

RELATION OF BLOOD PLATELETS TO PURPURA.—In many cases of purpura there is a reduction in the platelet count (*thrombocytopenic* purpura) which may be mild, *e.g.* down to 50,000, or severe, *e.g.*, down to 10,000 or even 1000 per c.mm.; in some cases, however, the platelet count remains normal (*athrombocytopenic* purpura). With low platelet counts though the coagulation time is normal, the clot that is formed is soft and friable, does nor retract well, and is doubtless a less satisfactory bung for damaged capillaries.[1] It is clear, however, that a low platelet count alone cannot be responsible for the initial *occurrence* of hæmorrhages. The injection of agar serum into an animal greatly reduces the platelet count, but does not produce hæmorrhages. The bleeding may cease in a severe case of purpura several days before the platelet count rises, and the platelets have been known to disappear clinically from the blood without hæmorrhages occurring. It is possible that in some cases of purpura enough platelets "stick" on to the wall of damaged capillaries (as they "stick" on to any injured endothelium, p. 148) to lower the count in the peripheral blood; alternatively the factors responsible for the capillary abnormality may be independently destroying the platelets. Thus in "hypersplenism" there is clear evidence of excessive platelet destruction, often associated with purpura (p. 229).

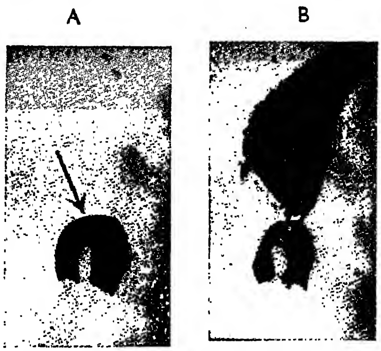


FIG. 89.—Response of Capillaries in Congenital Telangiectasis to Needle Puncture.

A. Greatly dilated capillary loop forming a telangiectasis; arrow shows needle track.
B. Extensive hæmorrhage from needle track which eventually had to be arrested by pressure. (After Macfarlane, *Quart. J. Med.*, 1941, *10.*)

Splenectomy cures about 70% of cases of severe thrombocytopenic purpura. The platelet count usually rises, but does not run parallel with the arrest of the hæmorrhage. It has been suggested that the spleen may in some unknown way be responsible for the capillary abnormality as well as for the thrombocytopenia.

Local bleeding in purpura can be controlled by firm pressure which closes the damaged capillaries and permits clotting to occur locally.

HÆMORRHAGIC TELANGIECTASIS.—This rare condition is an interesting example of a hæmorrhagic state unquestionably due entirely to a localized capillary abnormality. The essential lesion is the telangiectasis or group of

[1] Normal blood is allowed to clot round a glass rod and is then kept for one hour at 37° C.; when the clot adherent to the rod is removed, 40% of the original volume of the blood is left behind in the form of serum. When the experiment is repeated with blood from a case of thrombocytopenic purpura, the clot formed is soft and watery and when it is removed on the rod only a small volume of serum is left behind.

dilated capillaries in the skin or mucous membranes; these do not react to stimuli which affect normal capillaries. Profuse bleeding follows rupture of these vessels (Fig. 89).

DEVELOPMENT AND PROPERTIES OF THE RED BLOOD CORPUSCLES. NORMAL BLOOD STANDARDS

Development.[1]—IN THE EMBRYO.—In the early embryo, blood formation takes place first in the mesoderm of the yolk sac (the area vasculosa) and later in the body of the fœtus. The mesoderm consists originally of a syncytium or nucleated mass of protoplasm without cell outlines. The syncytium differentiates to give rise to a network of capillary vessels lined by endothelium and containing plasma (which is formed by liquefaction of the cytoplasm). Erythropoiesis takes place *intravascularly*; in places the endothelial cells proliferate and differentiate to form masses of nucleated hæmoglobin-bearing

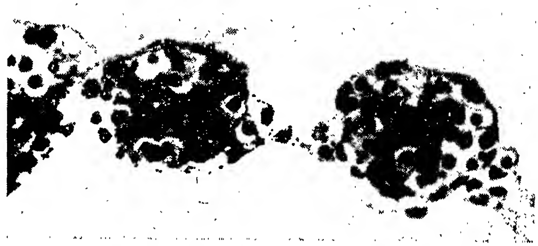


FIG. 90.—Erythropoiesis in Human Embryo. Blood vessels in yolk-sac full of nucleated erythroblasts. (Gilmour, *J. Path. Bact.*, 1940, *52*.)

cells which fill and distend the capillary lumen (Fig. 90). These cells become free and circulate in the blood stream and finally lose their nuclei to give rise to non-nucleated discs. The nucleated red cells disappear completely from the blood about the middle of fœtal life. After the third month of fœtal life the *spleen* and especially the *liver* are the most important sites of intravascular blood formation, nucleated red cells developing from the endothelial lining of the blood channels. About the middle of fœtal life the *bone marrow* begins to act as a blood-forming organ, the function becoming progressively more important as erythropoiesis in the liver decreases, so that the marrow is *normally the sole region where red cells are formed after birth.* (The white blood cells also develop mainly in the red marrow, the granulocytes being formed exclusively there.) Occasionally in adult life when the marrow cavity is nearly obliterated by sclerosis or other diseases, the spleen and liver again become important sites of blood formation.

Bone Marrow.—(1) DISTRIBUTION.—Bone marrow may be yellow or red. *Yellow* marrow consists of fat cells, blood vessels, and a minimal

[1] Blackfan, Diamond, and Leister, *Atlas of Blood in Children*, N.Y., 1944. Israëls, *Atlas of Bone Marrow Pathology*, London 1948. Gilmour, *J. Path. Bact.*, 1941, *52*, 25. Dacie, *J. clin. Path.*, 1949, *2*, 1. Leitner, *Bone, Marrow Biopsy*, London, 1949.

framework of *reticulum cells* and *fibres*[1]; in *red* marrow numerous *blood cells of all kinds and their precursors* (*erythroid* and *myeloid*[2]) are present too. At *birth* all the bones are filled throughout their length with highly cellular red marrow. With increasing age the marrow becomes more fatty, the process setting in first in the distal bones of the limbs (tarsus and carpus), then in the intermediate (tibia, fibula, radius, ulna), and finally in the proximal bones (femur, humerus) (Fig. 91). At the age of twenty all the marrow of the long bones is yellow except for the upper end of the femur and humerus. In the adult, red marrow persists mainly in the vertebræ, sternum, ribs, and bones of the skull and pelvis. There is obviously ample room in the long bones of the adult for considerable expansion of the red marrow. Children, in relation to their weight, have relatively more red marrow than adults

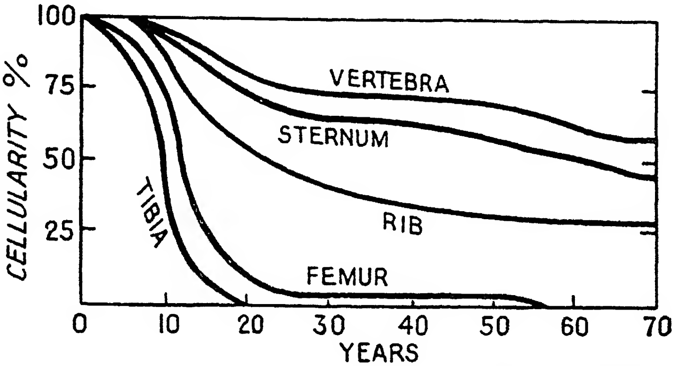


FIG. 91.—Changes in Red Bone Marrow with Age. (Whitby and Britton, *Disorders of the Blood*, 6th. edn., 1950, Churchill.)

The chart shows the decrease in cellularity of the red marrow of different bones with advancing age. The curve for tibia=also fibula, radius, and ulna. The curve for femur=also humerus; the curve for vertebræ=also pelvis. 100%=degree of cellularity at birth.

but little reserve space on which to draw if needed; in times of stress, *e.g.* in fibrosis of the red marrow or after severe hæmorrhage, blood formation may take place in other regions, *e.g.* liver and spleen. In adults, examination of samples of marrow aspirated from the sternum or iliac crest during life

[1] Reticulum cells are found in the adult in lymph nodes, spleen, bone marrow, and subcutaneous tissue. In the quiescent state they form a syncytium and are closely associated with reticulum fibres, also called *argyrophil* fibres because they reduce and are thus stained by silver salts. The framework of bone marrow, lymph nodes, and splenic pulp consists of reticulum cells and fibres. When appropriately stimulated the cells differentiate along various paths; as explained on p. 215, they give rise to the precursors of all varieties of white cells, *i.e.* myeloblast, lymphoblast, and monoblast. [The workers who believe that red cells also develop extravascularly (p. 163) regard the reticulum cell as the precursor of the red cell series too.] If stimulated by infection the reticulum cells proliferate and give rise to cells which wander in the tissues (and are then called macrophages, histiocytes, or wandering cells).

[2] *Hæmatopoiesis*=formation of blood cells; *leucopoiesis*=formation of white blood cells; *granulopoiesis*=formation of granular white cells; *erythropoiesis*=formation of red blood cells; *erythroid* cells=red cells and their precursors (also called collectively *the erythron*); *myeloid* cells=white blood cells and their precursors.

gives much information about the state of marrow activity. Post-mortem the sternum, a rib or a vertebra is examined when *diminished* marrow activity is suspected; the shaft of a long bone is generally studied when evidence of *extension* of hæmatopoiesis is sought.

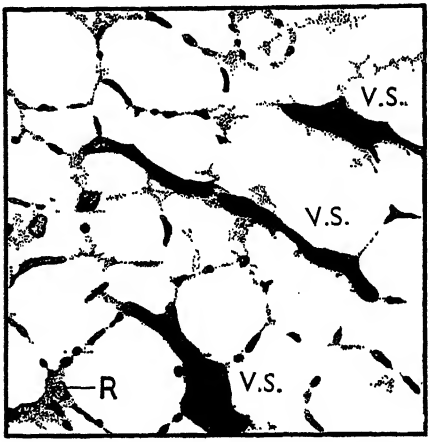


FIG. 92.—Sinusoids of Bone Marrow. Section of Hypoplastic Bone Marrow (Pigeon) injected with Indian Ink.

Three venous sinusoids (V.S.) are filled with the injection mass; extending between them are collapsed sinusoids marked by fine lines of particles of carbon between double rows of excessively thin endothelial cells. R=reticulum cell. (Doan, Cunningham, and Sabin

(2) VASCULAR ARRANGEMENT.—The nutrient artery of the bone breaks up into smaller branches, which lead to a network of intercommunicating *sinusoids*. These vessels are lined by a thin endothelium (like capillaries elsewhere) but when dilated have the capacity of large veins. Many of the sinusoids under normal conditions are completely collapsed and quite impermeable to blood; they can only be demonstrated when the cellular elements of the marrow have disappeared (*v. infra*). The injection of Indian ink into the nutrient artery then reveals these vessels as minute black streaks lined by very thin endothelial cells between every few fat cells (Fig. 92). Other vessels, however, are widely dilated and the blood flow through them is in consequence very slow. Some authorities believe that the collapsed sinusoids are the site of

---

MATURATION IN ERYTHROCYTE SERIES.

KEY TO PLATE II.

1. *Early normoblast* [*megaloblast*].
2. *Intermediate normoblast* [*erythroblast*].
3. *Late normoblast* [*normoblast*].
4. *Polychromatophilic erythrocytes* or *polychromatocytes:* young erythrocytes, still retaining some of the basophil staining reaction as well as the reddish hæmoglobin background.
5. *Punctate basophilic stippling*, with black or bluish granules in the cells. This is usually seen in severe anæmia and in toxic disturbances.
6. *Hypochromic erythrocytes:* cells often have only a ring of hæmoglobin at the periphery, and are usually small.
7. *Cabot rings:* remnants of nuclear membrane appearing as circles or figures of eight in the cells.
8. *Howell-Jolly bodies:* nuclear remains of varying size staining dark blue or black.
9. *Reticulocytes:* young erythrocytes, staining with a vital dye such as cresyl blue.
10. *Normal mature erythrocytes* [*normocytes*].
11. *Macrocytes* and *microcytes:* cells larger and smaller than normal.
12. *Poikilocytes:* cells of abnormal shape.
13. *Sickle cells:* in sickle cell anæmia the sickle shape is produced by depriving susceptible fresh cells of oxygen or exposing them to carbon dioxide.

Reproduced from Plate I, Blackfan, Diamond and Leister, *Atlas of the Blood*. The Commonwealth Fund, New York, 1944.

PLATE II

[*Facing page* 162

erythropoiesis and therefore call them *erythropoietic capillaries*. If this view is correct then *erythropoiesis throughout life is intravascular*. Most workers claim on morphological grounds that in man erythropoiesis in the red marrow is *extravascular*, the red cell precursors being the reticulum cells (p. 161); but they do not explain how the newly formed, non-motile red cells migrate to, and pass through, the intact endothelium of the marrow sinusoids. The site of erythropoiesis probably varies with the species (*infra*). (Leucopoiesis is always *extravascular* (p. 215).)

(3) Functions of the Red Marrow.—These are (i) development of red blood corpuscles (*infra*); (ii) development of granulocytes and to a less extent of monocytes and lymphocytes (p. 215); (iii) development of blood platelets (p. 155); (iv) destruction of red cells by macrophages which constitute part of the lining of the blood sinuses (and are therefore also called *littoral* cells (p. 187).

(4) Experimental Methods of Study.—Normal red marrow contains many types of cells which are so closely crowded together that a study of their inter-relationship is very difficult. To facilitate interpretation of the microscopic appearances the marrow is artificially simplified. If a *pigeon* is starved for 10–17 days all the cellular elements disappear from the red marrow, which now resembles normal yellow marrow. When feeding is resumed erythropoiesis soon recommences and can be studied satisfactorily. After a longer interval granulopoiesis begins again.

(5) Cytological Methods.—(i) Smears of living marrow and blood cells may be stained (vital staining) with Janus green and neutral red; the former stains the mitochondria and the latter the cell granules. (ii) Fixed marrow sections, marrow smears or blood films are stained with eosinate of methylene blue (Leishman's stain), to show nuclear details and the basophilia or eosinophilia of the cytoplasm (see Plate II). (iii) Living erythroid cells may also be stained with cresyl blue to demonstrate the presence of a *reticulum* in the cytoplasm (p. 165 and Plate II).

Stages of Erythropoiesis.—In certain species when red cell formation is actively taking place many of the marrow sinusoids which were previously patent become collapsed; the local blood flow is diminished, and this intensifies the anoxia which in turn favours erythropoiesis (cf. p. 194). The endothelial lining of the collapsed capillaries becomes swollen, differentiates and divides repeatedly to give rise to increasingly mature red cells which fill the sinus completely; the young cells lie externally near the lining of the sinus, and the more mature cells lie in the centre. The surface membranes of the immature nucleated red cells are "sticky," so that the cells adhere to one another, forming a solid mass. As the cells become mature and lose their nuclei the surface membrane is altered so that the corpuscles begin to separate one from the other. The plasma then percolates between the cells, which are finally swept out of the erythropoietic capillaries into the veins of the marrow, and so into the general circulation.

The account of erythropoiesis given above is based on studies of experimental animals; but as already mentioned erythropoiesis in man probably takes place extravascularly from the transformation of reticulum cells.

There is, unfortunately, no internationally accepted terminology for the intermediate forms between the endothelial cell and the mature erythrocyte. The current British terminology (which will be used in this book) and that

of Sabin, which is in general use in America, are compared in the Table below.

Cell stages.[1]	British Terminology.	American Terminology.	Staining of cytoplasm.	Mitosis.	Number per 100 nucleated erythroid cells in marrow.
I.	Proerythroblast.	Megaloblast.	Basophil.	+In states of stress.	5—
II.	Early normoblast.	Early erythroblast.	Basophil.	+	5—
III.	Intermediate normoblast.	Late erythroblast.	Polychromatophil.	+	20+
IV.	Late normoblast.	Normoblast.	Eosinophil.	—	70
	Reticulocyte.	Reticulocyte.			
	Erythrocyte.	Erythrocyte.			

[1] Davidson's terminology (cf. Fig. 110).

The British workers use the term *erythroblast* to include *all* nucleated red cells, normal and pathological; they restrict the term "megaloblast" to the characteristic nucleated red cells found in the marrow in cases of pernicious anæmia. Normal erythropoiesis is called *normoblastic*; the type of erythropoiesis found in pernicious anæmia is called *megaloblastic* (p. 196). Cell stages II, III, and IV in the red marrow of pernicious anæmia are called *early*, *intermediate*, and *late megaloblasts*.

The Table shows that most of the nucleated erythroid cells in the marrow are the later, more mature forms (intermediate and late normoblasts) which are normally being steadily transformed into erythrocytes; in times of stress the final process of maturation is speeded up leading to the rapid discharge of young red cells into the circulation.

Cytology of Erythroid Cells (Plate II and Fig. 93).—In general, developing marrow cells (myeloid or erythroid) show the following changes: the cells decrease in size and the cytoplasm becomes relatively more extensive; nucleoli disappear and the chromatin becomes progressively coarser; the cytoplasm becomes less basophilic and the pigment or granules characteristic of maturity appear. Developing red cells stained with Leishman's stain present the following appearance:

(1) Stage I, Proerythroblast (Megaloblast).—This early cell, which is derived from the endothelium, is large (diameter 15–20 μ); the cytoplasm is a deep violet-blue and there is a small crescent, showing paler staining, round the nucleus; the cell is devoid of hæmoglobin. The nucleus is large (12 μ) occupying about three-quarters of the cell volume and the chromatin forms a fine stippled reticulum and contains several nucleoli; it develops into the early normoblast. The proerythroblast only shows mitosis in states of stress.

(2) Stage II, Early Normoblast (Early Erythroblast).—This cell is somewhat smaller and shows active mitosis; the nucleoli have disappeared and the chromatin network is fine and shows a few nodes of condensation.

(3) Stage III, Intermediate Normoblast (Late Erythroblast).—The cell is still smaller (diameter 10–15 μ) and shows active mitosis; the resting nucleus shows further condensation of the chromatin; hæmoglobin is beginning to appear, its eosinophil staining giving the cytoplasm a polychromatic appearance.

(4) Stage IV, Late Normoblast (Normoblast).—This cell represents a maturation of the previous stage; mitosis has now ceased; the cell diameter is 7–10 μ; the nucleus is small, the condensed chromatin assuming a "cart-wheel" appearance and finally becoming uniformly deeply stained (*pyknotic*). The hæmoglobin accumulates to its maximum level, the cytoplasm giving an eosinophil reaction as a result. Pyknosis is a stage in the degeneration of the nucleus which breaks up and finally disappears owing to extrusion or lysis; a young red cell (reticulocyte) is thus formed.

Maturation of the erythroblasts thus involves a decrease in the size of the cell, increased condensation and finally pyknosis of the nucleus, accumulation of hæmoglobin, and a change in staining reaction of the cytoplasm from basophil via polychromatophil to eosinophil.

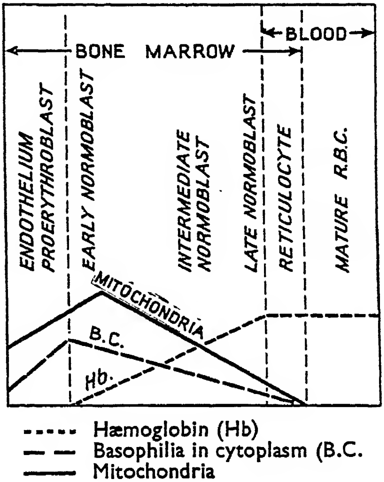


Fig. 93.—Chart showing Stages in Development of Red Blood Corpuscles.

(5) Reticulocyte. — The young red cell is so called because on vital staining with cresyl blue a network or reticulum is apparent in the cytoplasm in the form of a heavy wreath, or as clumps of small dots, or as a faint thread connecting two small nodes. (*All the nucleated precursors* of the reticulocyte (erythroblasts) also give this staining reaction.) The reticulum probably consists of remnants of the basophil cytoplasm of the immature cell (chemically the reticulum is made up of ribose-nucleic acid).[1] As the red cell ages, the reticulum disappears. In the new-born, 30–50% of the red cells in the circulation are reticulated; the number falls during the first week to less than 1%, at which level it remains throughout life. Their number is *increased whenever red cells are being rapidly manufactured.* For example, if hæmolytic poisons are injected intravenously to destroy the circulating red cells, the marrow proliferates, and numerous young red cells pass into the blood stream

[1] If red cells are stained with eosin and methylene blue the presence of the reticulum in the young cells (reticulocytes) leads to a diffuse bluish staining of the cell—*polychromatophilia.* In pathological states this stained basophil material is sometimes present in clumps which appear as discrete blue particles. This finding, known as *basophil punctation* (or *punctate basophilia*) (Plate II, 5), is especially obvious after poisoning with lead.

(Fig. 97) ; it is then found that 25–35% of the circulating red cells are reticulocytes. An increase in the reticulocyte count (*reticulocytosis*) is the first blood change noted when pernicious anæmia is treated with liver extract (p. 198).

(6) NORMAL RED CELL.[1]—This is a circular, non-nucleated biconcave disc ; it is markedly *elastic* and can undergo astonishing deformation when passing through narrow capillaries.

The *permeability* of the red cell membrane is fully discussed on p. 7. The membrane is freely permeable to Na^+, K^+ and Cl' ; in spite of this fact, however, the Na^+ and the Cl' content of the red cell is low though that of the plasma is high ; and the K^+ content of the red cell is high though that of the plasma is low. The red cell, using energy derived from glycolysis, " pumps " Na^+ and probably Cl' into the plasma ; secondarily the non-penetrating

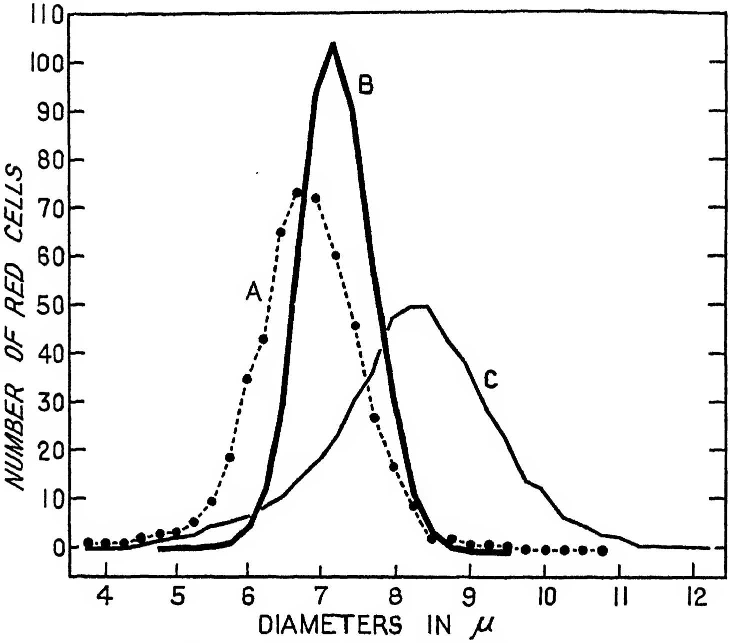


FIG. 94.—Diameter of Red Blood Corpuscles in Different Conditions. (Price-Jones).

A = Hæmorrhagic Anæmia (mean 6·85 μ).
B = Normal (mean 7·21 μ).
C = Pernicious Anæmia (mean 8·24 μ).

anions in the cell (protein and organic phosphate) retain the K^+. The composition of red cells is compared with that of plasma in Fig. 5. When cell metabolism ceases, as in cold-stored blood, the ions move between the cells and the plasma according to their concentration gradients (p. 8). When the CO_2 tension in blood rises, HCO_3' moves out of the cell into the plasma while Cl' moves in the reverse direction (chloride shift, p. 420).

Diameter.—The average *diameter* of the red cells in health is 7·3 μ, the

[1] Jacobs, *Ann. N.Y. Acad. Sci.*, 1950, *50*, 824 ; Waugh, *ibid.*, 835 ; Ponder, *ibid.*, 1947, *48*, 579.

range being 5·5–8·8 μ; in pernicious anæmia the average size is *greater* (about 8 μ), and the range is 4–12 μ. The distribution of size of red cells is usually set out graphically (Fig. 94) constituting a *Price-Jones curve.* When the H^+ ion concentration of the blood rises, the size of the corpuscles increases; this is presumably due to the migration of chloride from the plasma into the cells, raising their osmotic pressure and thus attracting water into their interior.

NORMAL RED CELL COUNT.[1]—The *average* red count in the adult male is 5·5 millions per c.mm.; in the female it is 4·8 millions. Counts which regularly fall below 4·5 million in the male and 4 million in the female should arouse suspicion. A brief period of vigorous exercise may increase the count, mainly owing to the passage of fluid out of the circulation (cf. p. 227). In new-born infants the red cell count is very high (6–7 million) (Fig. 122), but it falls rapidly during the next two weeks to a little above the average adult value.

Normal Blood Standards.—In describing the red cells it is necessary to give first (i) the *red blood count per c.mm.* (*supra*) and (ii) the *hæmoglobin content in g. per* 100 c.c. (p. 175). The following additional information is desirable, especially when considering the anæmias:

(iii) The *average diameter* (*and range of diameter*) *of the red cells* (p. 166 and Fig. 94). Cell diameter is usually (but not invariably) directly related to cell volume (cf. (v) *infra*).

(iv) *Relative volume of packed red cells.*—This is determined by centrifuging blood in a graduated tube (hæmatocrit) till the volume of the packed corpuscles which accumulate at the bottom of the tube becomes constant.[2] The volume of the packed red cells (hæmatocrit value) is normally 41–45% of that of the whole blood (cf. p. 10).

(v) *Mean corpuscular volume* is determined thus:

$$\frac{\text{Volume of packed red cells in c.c. per 1000 c.c. blood ((iv) \textit{supra})}}{\text{Red cells in millions per c.mm.}}$$

The answer is in cubic microns (μ^3 or c.μ.) and gives the dimensions of the cells in all planes and not only in the long axis as in diameter measurements ((iii) *supra*). The normal range is 78–94 μ^3. When the volume exceeds 94 μ^3 the cells are called *macrocytes*; when the volume is 78–94 μ^3 the cells are called *normocytes*; when the volume is less than 78 μ^3 the cells are called *microcytes*.

(vi) *Average corpuscular hæmoglobin content* (average amount of hæmoglobin per cell in micro-micrograms ($=10^{-12}$ g.) is determined thus:

$$\frac{\text{Hæmoglobin in grammes per litre of blood}}{\text{Red cells in millions per c.mm.}}$$

The average value is $29{\cdot}5\times10^{-12}$ g.$\pm2{\cdot}5$. This value is a more precise expression of the colour index. The *colour index* is the ratio $\frac{\text{Hæmoglobin \%}}{\text{Red Cell \%}}$ where 100% for the red cells is 5,000,000 red cells/c.mm. of blood and for

[1] The standard methods of counting red cells have an error of $\pm$15%; thus with a red cell count of 5 million the *technical* error of a *single* count may be as great as $\pm$0·75 million.

[2] The anticoagulant added to the blood must not alter the volume of the cells; heparin or a mixture of potassium and ammonium oxalates is used.

Fragility in Disease.—To understand the changes in fragility that take place in abnormal states we must consider in detail the sequence of events when red cells are placed in hypotonic saline. Fig. 96 gives representative values for the diameter (7·6 μ), thickness (1·9 μ), volume (86 μ^3), and surface area (138 μ^2) of a normal red cell. When it is placed in hypotonic saline, fluid flows into the cell and it becomes more spherical, the diameter diminishing somewhat and the thickness increasing considerably. If the volume and surface area changes are considered the results are somewhat unexpected. When a biconcave disc-like cell is converted into a sphere of equal volume its surface area is reduced; therefore, as the red cells become more spherical in hypotonic saline they can accommodate the additional fluid at smaller surface areas. Finally the cell becomes fully spherical and attains its

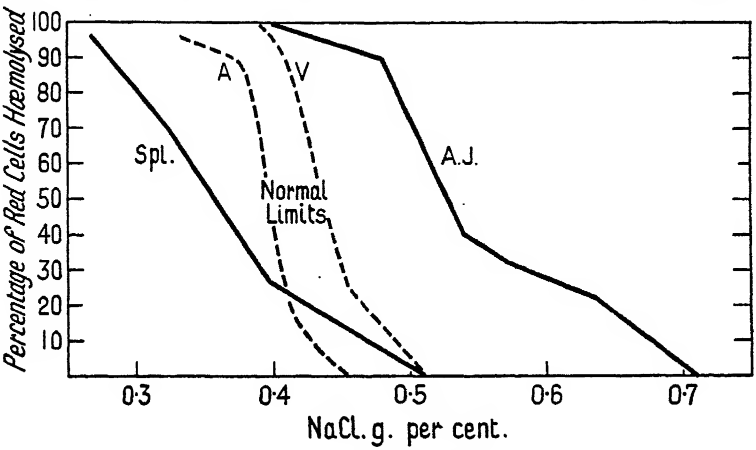


Fig. 95.—Quantitative Estimation of Fragility of Red Blood Corpuscles.

A–V = normal limits; A = arterial blood; V = venous blood; Spl = after splenectomy; A.J. = acholuric jaundice (congenital hæmolytic jaundice) (cf. Fig. 96 and p. 229). (After Whitby and Hynes, *J. Path. Bact.*, 1935, *40*.)
Abscissa = concentration of NaCl employed for diluting corpuscles.

original surface area with a greatly increased volume (*e.g.* 152 μ^3 instead of 86 μ^3, or +77%). The membrane *cannot stretch*, so that if there is further inflow of fluid beyond this point, rupture takes place and hæmolysis occurs.

The behaviour of a red cell in hypotonic saline depends then on the initial ratio of surface area to volume and not on the absolute size of the cell. The greater this ratio the larger the additional volume that can be accommodated in a sphere of equal surface area, and conversely.

In some anæmias when one dimension of a red cell is altered, the other alters proportionately; in such cases fragility is unaffected. But in other anæmias the mean cell diameter is increased, while the thickness is decreased (Fig. 96). In idiopathic hypochromic anæmia cell volume, diameter and thickness are usually all decreased; sometimes however the decrease in thickness is relatively more marked and in extreme cases this gives the stained cell an appearance which has suggested the name of "target cell." These thin cells can swell by +128% (Fig. 96, 1) before rupture occurs; they

can thus withstand the action of a more hypotonic saline and are said to show diminished fragility. Hæmolysis may begin only at 0·36% saline and may not be complete until 0·24% solutions are used. Conversely, in congenital hæmolytic jaundice (acholuric jaundice) (p. 229) the cells have a smaller diameter (*e.g.* 6·4 μ, but are fatter (thickness 2·4 μ), *i.e.* they are *spherocytic* (*spherocytes*) (Fig. 96, 3). The ratio of surface area to volume is subnormal. An increase in volume of only 30% (from 77 μ^3 to 101 μ^3) converts them into fully distended spheres at which stage rupture occurs. They are said to show excessive fragility, *i.e.* they rupture in the higher (more concentrated) saline solutions that do not cause the rupture of normal cells ; thus hæmolysis may begin in 0·7% saline and be complete at 0·45% (Fig. 95). The term high or

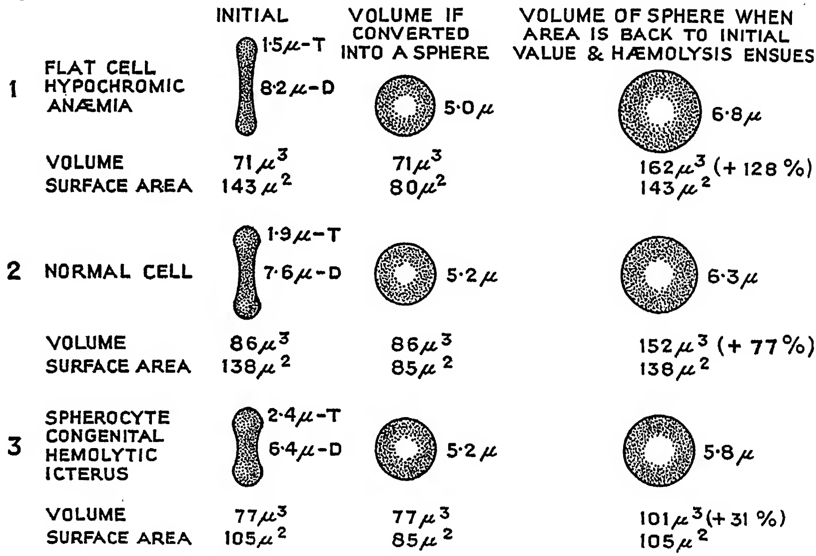


FIG. 96.—Relation of Red Cell Volume to Surface Area in Different Conditions. (Haden, *Symposium on Blood*, 1939.)
T = thickness, D = diameter, of red cell.

low fragility gives no information about the state of the cell membrane. The membrane is not stretchable and does not show varying resistance to distension in health or disease. The fragility of red cells in hypotonic saline is related to *cell shape* ; the practical rule is that flat long cells withstand weak NaCl solutions, fat short cells (spherocytes) rupture in stronger NaCl solutions. Fragility results should always be interpreted in this manner.

Cells of venous blood are slightly more fragile than cells of arterial blood in the same normal subject (Fig. 95) owing to a slight spherocytosis of the venous cells.

The effects of *splenectomy* on red cell fragility in normal subjects and disease conditions are dealt with on pp. 229, 231 (cf. Fig. 95).

ACTION OF HÆMOLYSINS.[1]—Hæmolytic sera can be prepared by injecting red cells repeatedly into another species. Thus, if guinea-pig red cells are

[1] Dameshek, *Ann. N.Y. Acad. Sci.*, 1947, *48*, 685.

injected into rabbits, the serum of the latter will contain hæmolysins to guinea-pig cells. If such hæmolytic sera are injected into guinea-pigs the first effect on the blood is increased spherocytosis of the cells as shown by a decrease in diameter, an increase in thickness, and increased fragility in hypotonic saline. Subsequently hæmolysis takes place, its degree (*i.e.* the extent of fall in red cell count) depending on the amount of initial spherocytosis. The red marrow responds by increased activity, turning out young

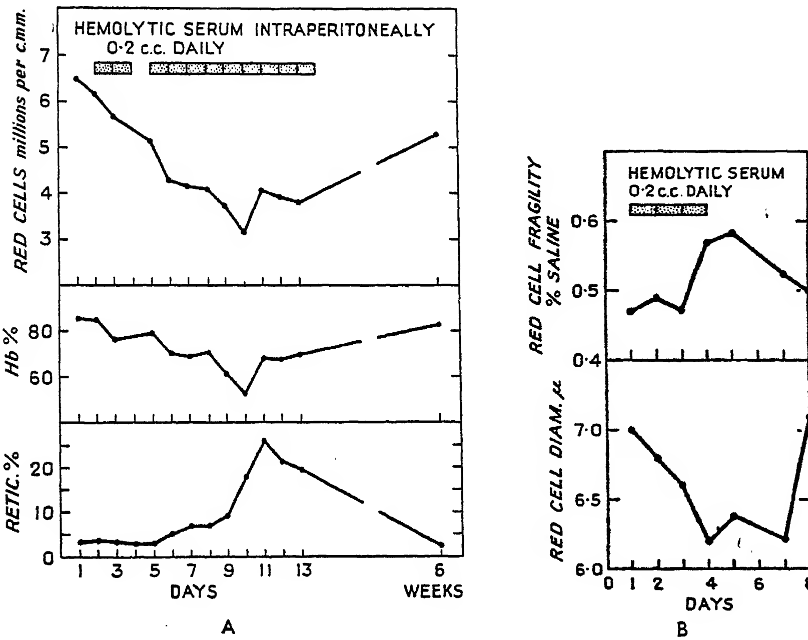


FIG. 97.—Action of Hæmolytic Sera on Red Blood Cells.

Experiments on guinea-pigs. Inject daily 0·2 c.c. of serum containing hæmolysin against guinea-pig red cells.

A. Note development of progressive anæmia (fall of hæmoglobin and red cell count) and active response of red bone marrow as indicated by the increase in the reticulocyte count. Spherocytosis develops with the progress of the anæmia.

B. Note the progressive decrease in red cell diameter (spherocytosis) which is proportional to the increased fragility of the cells in saline solutions. (Dameshek and Schwartz, *Amer. J. med. Sci.*, 1938, *196*.)

red cells (reticulocytes) in large numbers and finally restoring the blood to normal (Fig. 97, A and B).

Hæmolytic Anæmia in man may be divided into five groups :

(i) The cells are defective in structure and, therefore, especially prone to succumb to the normal hæmolytic processes of the body : *e.g.* pernicious anæmia (p. 196), congenital hæmolytic jaundice (p. 229), sickle cell anæmia and Mediterranean anæmia.

(ii) The red cells are normal in structure but are destroyed excessively, together with leucocytes and platelets, by an abnormally acting spleen, *e.g.* in the panocytopenia of so-called hypersplenism (p. 230).

(iii) Resulting from the action of known simple or complex lysins introduced from without: arseniuretted hydrogen, snake venom, streptococcal hæmolysin.

(iv) Resulting from the action of biological hæmolysins of uncertain nature, usually elaborated by the patient himself; paroxysmal hæmoglobinuria (p. 177), nocturnal hæmoglobinuria, erythroblastosis fœtalis (p. 180).

(v) Atypical obscure hæmolytic anæmia. Bloods of the fourth group and some of the fifth group give positive Coombs tests.

Coombs test.—Autohæmolysins of the type described are globulin in nature, adsorb strongly on to the red cell and unlike the normal globulins of the blood plasma cannot be removed by washing with saline. Hence when the affected cells are exposed to the serum of a rabbit immunized against human globulin, they are clumped, while normal washed cells are unaffected; this is the *direct* test. If the hæmolysin is free in the plasma of the patient, it will adsorb on to normal compatible red cells (if brought in contact with them). These will probably not be lysed but even after washing with saline will clump on the addition of anti-human-globulin rabbit serum; this is the *indirect* test. The tests have a diagnostic value in erythroblastosis fœtalis (p. 180).

Sedimentation Rate of Red Blood Corpuscles.[1]—If blood (kept fluid by means of an anticoagulant) is allowed to stand in a narrow tube the corpuscles settle progressively to the bottom while the plasma rises to the top. The rate at which this takes place is very constant in health and is known as the *sedimentation rate.* The corpuscles settle because their density is greater than that of the plasma. Settling is resisted by the viscosity of the medium; the retarding force so exerted varies directly with the surface area of the falling particles. Cells allowed to stand outside the body tend to aggregate to form *rouleaux* (cells piled one on top of the other, a condition which must be distinguished from agglutination (p. 176) where the cells are irreversibly clumped together and cannot be separated); as they do so, the ratio of mass to surface area in the enlarged particles increases, and as a result they sink with greater rapidity. Thus if a red corpuscle is regarded as a short cylinder of diameter 8 μ and height 2 μ, its surface area would be 150 sq. μ; ten *discrete* corpuscles would have a surface area of 1500 sq. μ. If, on the other hand, these corpuscles were piled on top of one another to build a rouleau the total free surface exposed to the plasma is only 600 sq. μ with corresponding speeding up of sedimentation. Fundamentally, therefore, the rate of sedimentation varies with the speed of rouleau formation. The factors controlling this process outside the body are not fully known, but the most important is the *composition of the plasma*: *fibrinogen,* other *globulins,* and certain *products of tissue destruction* increase the sedimentation rate; *albumin* reduces it.[2] Fibrinogen is the principal single factor: the sedimentation rate of defibrinated blood or of the blood of patients in whom the fibrinogen is absent is extremely slow; this point is well illustrated in Fig. 98. Conversely, one finds clinically that the sedimentation rate is faster whenever the fibrinogen level is raised.

[1] Ham and Curtis, *Medicine,* 1938, *17,* 447. [E.S.R.=erythrocyte sedimentation rate.]

[2] With higher red cell counts the sedimentation rate is normally slowed down, and this has to be allowed for when examination is being made of blood with abnormally high or low counts.

To determine the sedimentation rate, a long, narrow graduated tube is filled with citrated blood to the 10 cm. mark and kept at a temperature of 22–27° C. At the end of *one hour* the upper level of the red cells is read and the height in mm. of the *column of clear plasma* is noted. Normally the height of this column is 0–6·5 mm. in 90% of normal males and 0–12 mm. in normal females; the more rapid sedimentation in females is due to their lower red cell count. The average for males is 4 and for females 8 mm. The sedimentation rate is speeded up in pregnancy from the third month up to parturition. In pathological states the sedimentation rate may exceed 50 mm.

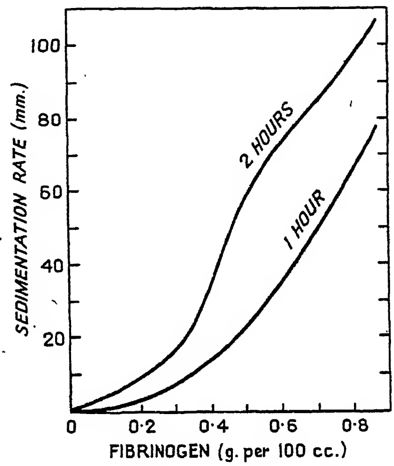


FIG. 98—Effect of Plasma Fibrinogen Level on Sedimentation Rate of Red Blood Cells.

Observations on boy with congenital absence of fibrinogen.
Ordinate: sedimentation rate in mm. (expressed as height of column of clear plasma) read at end of 2 hours (upper curve) and at end of 1 hour (lower curve).
A series of fibrinogen concentrations was prepared by mixing varying proportions of the patient's plasma with normal plasma. There is a negligible sedimentation rate in the absence of fibrinogen; note the progressive increase in the sedimentation rate with increasing fibrinogen concentration. (Redrawn from Oakley, *Lancet*, 1938, ii, 312.)

The sedimentation rate is usually unaffected in diseases without inflammation or tissue destruction. It is grossly increased in conditions associated with *inflammation* or *tissue destruction*, such as acute or chronic infections (local or general) of any severity, or severe trauma. This increase is usually associated with an elevated fibrinogen level and occasionally a rise in globulin. The presence of an increased sedimentation rate suggests organic disease, even in the absence of other signs; the finding, however, is quite *non-specific* and gives no indication of the nature of the disease. The main value of sedimentation rate determinations is in assessing prognosis or in judging progress or the effects of treatment. A rising sedimentation rate suggests worsening of the condition or heralds the onset of complications. In tuberculous infections and also in rheumatic fever, the sedimentation rate is widely employed to judge the course of the disease.[1]

The sedimentation rate is increased in *nephrosis*. In this condition the serum albumin is considerably reduced; the serum globulin is generally increased when there is pronounced lipæmia; (the lipides in the serum are attached to the globulins).

Hæmoglobin.[2]—Hæmoglobin consists of the protein *globin* united with the pigment *hæm*. Hæm is an *iron-containing porphyrin* known as *iron-protoporphyrin IX*. The porphyrin nucleus consists essentially of four *pyrrol* rings

[1] For general theoretical discussion, see Fåhræus, *Lancet*, 1939, ii, 630.

[2] Rimington, *Lancet*, 1951, ii, 551. Dobriner and Rhoads, *Physiol. Rev.*, 1940, *20*, 416. Theorell (Hæmoproteins), *Recent Advances in Enzymology*, 1947, 7, 265. Granick and Gilder (Tetrapyrrols) *ibid.*, 305.

joined together by four *methine* $\left(=\underset{\mathrm{H}}{\underset{|}{\mathrm{C}}}-\right)$ " bridges "; the porphyrins are thus *tetrapyrrols*. The " skeleton " of the formula of hæm is shown in Fig. 99; the pyrrol rings are numbered I, II, III, IV; the carbon atoms of the methine bridges are labelled α, β, γ, δ; the positions to which side chains are attached are numbered 1–8. The side chains at the respective positions are: 1, methyl ($-CH_3$); 2, vinyl ($-CH{=}CH_2$); 3, methyl; 4, vinyl; 5, methyl; 6, propionic acid ($-CH_2.CH_2.COOH$); 7, propionic acid; 8, methyl. Thus side chains 1, 3, 5, and 8 are methyl; 2 and 4 are vinyl; 6 and 7 are propionic acid.

The iron in hæm is in the *ferrous* (Fe^{++}) form. The iron is attached to the N of each pyrrol ring and to the N of the iminazol group in the associated globin; a " bond " is available for *loose union with* O_2 (in *oxyhæmoglobin*)

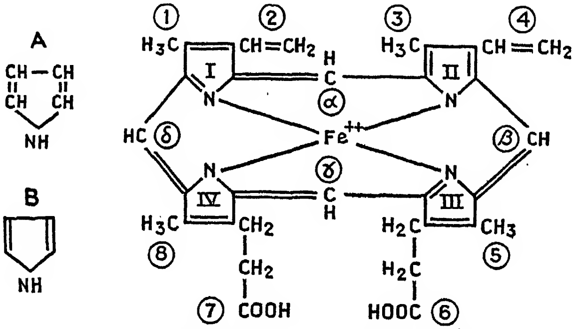


FIG. 99.—Chemistry of Hæm (Iron-Protoporphyrin IX).

A. Pyrrol ring in full.
B. Pyrrol ring, conventional outline.
Hæm: Pyrrol rings are numbered I, II, III, IV. C atoms of methine bridge are labelled α, β, γ, δ. Side chains are numbered 1–8.

or *CO* (in *carboxyhæmoglobin*). In *reduced* hæmoglobin this place is occupied by an OH group.

The hæm is further attached to the globin by a link between each of the two propionic acid side chains and basic groupings in the globin.

The molecular weight of hæmoglobin is 68,000. When treated with alkali and a reducing agent, a substance called *hæmochromogen* is formed with a molecular weight of 17,000. Hæmoglobin may thus represent the product of condensation of *four* hæmochromogen molecules.

The property which hæmoglobin possesses of combining loosely with oxygen and other gases depends on the presence of *ferrous* iron. Each atom of Fe^{++} combines with one molecule of O_2 or CO.

When reduced or oxygenated hæmoglobin is treated with an oxidizing agent, the Fe^{++} is oxidized to ferric iron (Fe^{+++}); the sixth bond is attached to OH. The compound is called *methæmoglobin*; it cannot unite reversibly with gaseous oxygen; the O of the attached OH is not given off in a vacuum. Reduced hæmoglobin is commonly represented as Hb; oxyhæmoglobin as HbO_2; methæmoglobin may be represented as HbOH.

Normal Blood Hæmoglobi .—In males the average hæmoglobin content of blood is 15·8 g. per 100 c.c.; in females 13·7 g. The average, irrespective of sex, is 14·5 g. In 90% of normal males the range is 14–18 g. and in females 12–15·5 g. At birth the hæmoglobin concentration is 23 g–%; by the end of the third month it has fallen to well below normal (average 10·5 g.) (cf. p. 207); recovery then gradually takes place and at the end of the first year the hæmoglobin reaches about 12·5 g–% (Fig. 100). 1 *g. of Hb when fully saturated combines with* 1·34 *c.c. oxygen,* so that the hæmoglobin concentration is an index of the *oxygen-carrying power* of the blood.[1]

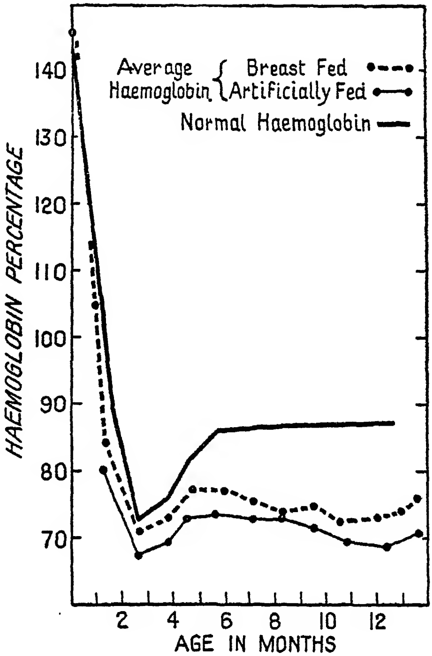


FIG. 100.—Changes in Hæmoglobin Concentration during First Year of Life. (Mackay, *Arch. Dis. Child.*, 1933.)

Hæmoglobin (on ordinate) = Haldane Scale. Normal hæmoglobin = values on a satisfactory diet (*infants receiving added iron*). The lower average values (both for breast-fed and artificially-fed babies) are due to inadequate nutrition.

FUNCTIONS OF HÆMOGLOBIN.—(1) It is essential for *oxygen carriage* (p. 409).

(2) It plays an essential part in the *transport of* CO_2 (p. 414) and in the regulation of *blood reaction* (p. 93).

The inclusion of hæmoglobin *in the corpuscles* gives certain advantages. If it were dissolved in the plasma it would increase its *viscosity,* and raise its *osmotic pressure* by about 100 mm. Hg, and so entirely derange the mechanism of water interchange between the capillaries and the tissue spaces (p. 18). Further, freed hæmoglobin is excreted by the *kidney* and taken up and *destroyed* by the reticulo-endothelial system (p. 187). Certain important *shifts of*

[1] Clinically hæmoglobin concentrations are not usually expressed as above, in g. per 100 c.c. blood, but in relation to an arbitrary normal standard called "100% hæmoglobin." The actual concentration of hæmoglobin to which this corresponds varies with the author; thus for Haldane "100% hæmoglobin" meant 13·8 g. per 100 cc.; 50% hæmoglobin thus meant $\frac{1}{2} \times 13{\cdot}8$ g. $= 6{\cdot}9$ g. per 100 cc. The Haldane standard has recently been re-examined and it has been found that 100% really corresponds not to 13·8 but to 14·8 g. of hæmoglobin per 100 cc. Previous conversions of "Haldane estimations" to g. of hæmoglobin must all be altered in the light of this finding. On the Sahli scale, 100% hæmoglobin means 27·3 g. per 100 cc. The expression "per cent. hæmoglobin" is therefore meaningless unless the standard adopted is known; in any case it is a most misleading and undesirable form of terminology and ought to be abandoned. At the moment this advice is counsel of perfection, as the use of the term is too deeply ingrained. In this book the Haldane standard is probably referred to when the expression "hæmoglobin per cent." is mentioned.

ions (Cl′ and HCO_3' especially) between plasma and red cells depend on the fact that hæmoglobin is enclosed in a cell bounded by a membrane with characteristic properties (p. 166).

Factors concerned in *hæmoglobin formation* are summarized on p. 194.

BLOOD GROUPS

Blood Groups.[1]—Three main groups of "factors" are present in human blood cells which enable the cells of different individuals to be differentiated; these factors have great clinical, medico-legal, and genetical interest. The groups are:

(i) A, B, and O.
(ii) Rh (Rhesus) factors.
(iii) M and N.

Classical (A, B, O) Blood Groups.—Human beings can be divided into four main groups according to the presence (or absence) in their red cells (and in certain tissue cells) of the substances called A, B, and O: 42% contain substance A; 9% contain substance B; 3% contain substance AB; 46% contain substance O. The percentage distribution given above is for Western European peoples.[2] The groups are correspondingly called group A, group B, group AB, and group O. More refined analysis shows that substance A can be subdivided into two sub-groups called A_1 and A_2; A_1 includes 75% of all group A, A_2 forms 25%. Group AB is similarly divided into A_1B and A_2B.

A and B are called *group specific* substances and chemically are *polysaccharides*; they are also *agglutinogens, i.e.* in the presence of a suitable antibody called an *agglutinin*, agglutination or clumping of the red cells occurs. The agglutinin acting on agglutinogen A is called α or anti-A; the agglutinin acting on agglutinogen B is called β or anti-B. [These agglutinins are also called *isohæmagglutinins.*] Group specific substance O does not normally act as an agglutinogen and there is no corresponding agglutinin[3]; group O cells are not agglutinated by agglutinins α or β.

The so-called Landsteiner's law states the following: if an agglutinogen is present in the red cells of a blood, the corresponding agglutinin must be absent from the plasma; if the agglutinogen is absent the corresponding agglutinin must be present. The first part of this law is a logical outcome of the situation, for if both agglutinogen and agglutinin were present the cells would be agglutinated. The second part is a fact but not a necessary sequence, for when the agglutinogen is absent from blood the agglutinin might well be absent too. In fact absence of Rh agglutinogen is not normally accompanied by presence in the plasma of the anti-Rh agglutinin; similarly absence of M or N substance is *not* accompanied by the presence of anti-M or anti-N in the plasma. It is only in the case of the A, B, O blood groups that absence of A is associated with the presence of anti-A (α) and the absence of B with the presence of anti-B (β). The finding is strange and at present inexplicable.

[1] Wiener, *Blood Groups*, Springfield, 3rd edn., 1943. Symposium, *Ann. N.Y. Acad. Sci.*, 1946, *46*, 883–992. Race and Sanger, *Blood Groups in Man*, Oxford, 1950.

[2] Some of the Eastern European peoples show a higher proportion (up to 20 per cent.) of Group B. Pure American Indians belong almost exclusively to Group O.

[3] Very rarely, after repeated transfusions, group O cells evoke an irregular agglutinin (anti-O) in susceptible recipients.

If the *serum* of the four classical groups is examined, the agglutinins are found to be distributed as follows: Group A contains β; Group B contains α; Group AB contains no agglutinin; Group O contains α and β. The full description of the four groups taking into account both agglutinogens and agglutinins would be, therefore: $A\beta$; $B\alpha$; AB; $O\alpha\beta$. The agglutinin α is not a single substance but can be subdivided into α_1 and α proper: α_1 only agglutinates A_1; α proper agglutinates both A_1 and A_2. The relative amounts of α proper and α_1 vary in the sera of Groups B and O.

To determine an individual's classical blood group a saline suspension of his red cells is mixed on a slide (i) with a test serum containing α, and (ii) with a serum containing β. The results are diagnostic as shown by the following table.

Cells.	Serum from Group A Subject. Agglutinin β [anti-B].	Serum from Group B Subject. Agglutinin α [anti-A].
A . .	−	+
B . .	+	−
AB . .	+	+
O . .	−	−

+ = agglutination. − = no agglutination.

When no agglutination occurs the red cells remain separate and evenly distributed; when agglutination occurs the cells are massed together in *clumps* and lose their outline. With high titre (*i.e.* powerful) agglutinins the cells are massed into a few large clumps; with weaker agglutinins more numerous but smaller clumps are formed.[1]

[1] The *agglutinogens* A and B first appear in the sixth week of fœtal life; their concentration at birth is one-fifth the adult level and it progressively rises during puberty and adolescence. Group specific substances A and B are not limited to the red cells but are found in many organs: salivary glands and pancreas ++; kidney, liver, and lungs +; testis. Being water-soluble they appear in the body fluids and in the following relative concentrations: saliva and semen, 600; amniotic fluid, 175; [*red cells*, 8–32]; tears, 5; urine, 3; cerebrospinal fluid, 0.

Only 50% of new-born infants have demonstrable agglutinin, which has simply filtered across the placenta from the mother. The specific agglutinins appear at 10 days, rise to a peak at 10 years, and then decline. At all ages there are marked variations in agglutinin content in different individuals: the range for the "titre" (*i.e.* concentration) of agglutinin α is 8–2048 (most at 128); for β, 2–1024 (most at 32). The figures quoted indicate the *extent to which the serum can be diluted before losing its agglutinating potency*; a high titre represents high agglutinin activity. The specific agglutinins act best at *low* temperatures and against well-diluted cells; with weak sera and high cell concentrations the cells may mop up the agglutinin without being agglutinated.

Non-specific agglutinins may occur clinically which act in the cold (at 0°–5° C.) and not at body temperature; a person's *cold* agglutinins may agglutinate his own red cells. In some cases the concentration of the cold agglutinins is sufficiently high to agglutinate the subject's own cells at room temperature. Exposure of the limbs to cold may then lead to intravascular hæmoglobinuria (*paroxysmal hæmoglobinuria*).

Relation of A, B, O Blood Groups to Blood Transfusions.—The effects of transfusing the blood of any group into the circulation of a member of another group can readily be worked out if the following additional points are borne in mind. The serum agglutinins of the donor can usually be ignored (if their potency is not too high) as they are sufficiently diluted by the much larger volume of the plasma of the recipient to produce no ill effects [1] and they are also neutralized by soluble agglutinogens which are found free in the recipient's body fluids. It follows, therefore, that the agglutinogens of the cells of the recipient need not usually be considered. Account need only be taken of the effect of the *serum agglutinins of the recipient on the cells (agglutinogens) of the donor.* The following table indicates the effects that would be produced by transfusing cells of any group into a recipient of any group (the sign + indicates agglutination of the cells and incompatibility; the sign — indicates no agglutination, and therefore compatibility):

Recipient (Agglutinins in serum in italics).	Red Cells of Donor.				Percentage of individuals in each group.
	AB	A	B	O	
AB+ *no agglutinin (o)* .	—	—	—	—	AB 3
A+ *agglutinin β* . .	+	—	+	—	A 42
B+ *agglutinin α* . .	+	+	—	—	B 9
O+ *agglutinins αβ* .	+	+	+	—	O 46

The following important conclusions can be drawn from the preceding table. Group A and Group B can only safely receive red cells from Group O and their own group. Group AB contains no agglutinins in the serum and has therefore been called the *universal recipients.* The cells of Group O contain no agglutinogen and are not agglutinated by the members of any group; Group O have therefore been called *universal donors.* The classical terms, universal donor and recipient, are however no longer valid as they ignore the complications produced by the existence of the Rh factors. It should be emphasized that the only safe method of determining compatibility is to test *directly* the serum *of the recipient against the donor's corpuscles*; this should, whenever possible, be carried out in practice. If carried out as described on p. 181, it will safeguard against Rh as well as against ABO incompatibility.

Rh (Rhesus) Blood Groups.[2]—The recently discovered Rh groups are of outstanding clinical importance. The original discovery was made as follows: red cells of the rhesus monkey were injected into rabbits; the

[1] There are, however, exceptions to this statement, and minor degrees of agglutination of the recipient's corpuscles may occur.

[2] Mollison, Mourant and Race, *Rh Blood Groups and their Clinical Effects*, British Med. Res. Council Sp. Rep. No. 19, 1948.

rabbit responds to the presence of an antigen in these cells by forming an antibody which agglutinates rhesus red cells. The surprising observation was then made that if the immunized rabbit's serum is tested against *human* red cells, agglutination occurs in 85% of white men; these people are called Rh+ (positive) and their serum contains no Rh antibody. No agglutination occurs in 15%; these are called Rh— (negative) and their serum also contains no Rh antibody except in the circumstances explained below.

There are several varieties of Rh antigen and of Rh antibody; the commonest Rh antigen is called D, and its antibody is called anti-D. Blood group antigens are the result of the action of genes which are present in the chromosomes. The gene corresponding to the antigen D is also called D; when D is absent from the chromosome, its place is occupied by the alternate form (*allelomorph*) called *d*. A Rh gene is inherited from both the father and the mother. If gene D is carried by both sperm and ovum the resulting gene composition (*genotype*) of the offspring is DD; if the gametes carry D and d respectively the result is Dd; if both gametes carry d, the result is dd. DD (called homozygous) and Dd (called heterozygous) are both Rh+; dd (homozygous) is Rh—. Of the 85% Rh+ English people examined, 35% were DD and 48% were Dd; the remaining 2% of the Rh+ people had some other genotype containing D. In the case of a *homozygous* father of genotype DD, all the sperm contain D; with a *heterozygous* father of genotype Dd, half the sperm contain D and half d. These factors are of importance in relation to the genotype of the child and the likelihood of development of hæmolytic disease (p. 180).[1]

If red cells containing D are repeatedly injected experimentally into Rh— subjects, anti-D is formed. The *first formed* anti-D is called *saline agglutinin* or *complete antibody*; *in vitro* it agglutinates cells containing D when they are suspended in saline or in an albumin solution. The *later formed* anti-D is called *albumin agglutinin*; *in vitro* it agglutinates cells containing D when they are suspended in an albumin solution but *not* in saline. But though the addition of "late" anti-D serum to D cells suspended in saline does not agglutinate them it modifies their properties in anohter way: the D cells so treated can no longer be agglutinated in saline by "early" anti-D. "Late" anti-D (albumin agglutinin) is thus called *blocking antibody* or *incomplete antibody*. Anti-D is formed in response to *intramuscular* as well as intravenous injections of blood; the volume of blood injected is unimportant: as little as 0·05 c.c. of blood may lead to striking anti-D formation. Variations in individual responsiveness are, however, very great;

[1] A fuller statement of the Rh groups is as follows: Human red cells may carry certain of the agglutinogens C, c; D, d; E, e; as explained below.

Three of these are always present at a time and the *presence of one of a pair excludes the other*: thus Cde and cDE are possible: cdD is not. It follows that eight possible arrangements of genes may be carried by the sperm and eight by the ovum giving rise to thirty-two genotypes in all. Of these, only five are common: CDe.CDe; CDe.cDE; CDe.cde; cDE.cde; and cde.cde. cDE.cDE is rare and the other twenty-six, very rare. Any of the six agglutinogens C, c; D, d; E, e; might theoretically give rise to agglutinins, but practically all are very feebly antigenic except D which is strongly antigenic and readily evokes anti-D. For this reason CDe.Cde; CDe.cDE; cDE.cDE; CDe.cde and cDE.cde may be grouped together as the Rh+ group (or group D) evoking anti-D in any blood not containing the agglutinogen D, which for practical purposes narrows down to cde.cde, *i.e.* the Rh— group or group d. Interaction between these groups covers 98% of all cases of erythroblastosis fœtalis and Rh blood transfusion reactions.

thus after a single transfusion of Rh+ blood only 50% of recipients develop anti-D.

The serum of Rh− subjects contains no anti-D except in two circumstances: (i) following a transfusion or injection of Rh+ blood; (ii) in women who have undergone one or more pregnancies with a Rh+ fœtus (*infra*).

Rh Factor and Haemolytic Disease.—The child of a Rh− mother (genotype dd) and a Rh+ father (genotype DD) must be Rh+ (Dd); if the Rh+ father is Dd the offspring may be Rh+ (Dd) or Rh− (dd). If the mother is Rh− and the fœtus is Rh+, serious complications may occur: Cells containing D pass across the placenta from the fœtus to the mother; the latter responds by forming anti-D which returns to the fœtal circulation and tends to destroy the fœtal red cells. The degree of damage done to the fœtus depends on the magnitude of the maternal anti-D response and the ability of maternal Rh agglutinins to cross the placenta. Generally no harm is done during the first pregnancy; but serious results may occur in the second or later pregnancies depending on the degree of sensitivity of the mother. If the mother has been immunized previously by a Rh+ blood transfusion at any time, *even in childhood*, a dangerously high response may occur during a first pregnancy. But it should be emphasized that in most cases, agglutinins are *not* formed and the great majority of matings between a Rh− mother and a Rh+ father result in normal offspring.

Effects of anti-D on the Fœtus.—The changes in the fœtus may be termed *hæmolytic disease* (not to be confused with *hæmorrhagic* states due to vitamin-K deficiency, p. 153) because they are due to the destruction of the red cells by maternal anti-D. The following are the chief clinical syndromes:

(i) *Hydrops Fœtalis:* the fœtus is grossly œdematous; it either dies *in utero*, or if born prematurely or at term, it dies within a few hours.

(ii) *Icterus Gravis Neonatorum* (*Congenital Anæmia of the New-born*).—The infant is born at term; it is jaundiced (hæmolytic jaundice, p. 190) or becomes so within 24 hours. There may be no anæmia at birth because the excessive red cell destruction is more or less compensated for by an intense normoblastic response of the marrow, associated with a high reticulocyte count and the presence of many nucleated red cells in the circulating blood (*erythroblastæmia*; *erythroblastosis fœtalis*). The rate of red cell destruction by anti-D is maximal at birth and so anæmia may develop in the first few days. There may be severe neurological lesions involving the basal ganglia especially; they secondarily become stained bright yellow with bile pigment (*kernicterus*) (p. 660); the liver may also be damaged. Free anti-D (derived from the mother) is present in the infant's blood for at least one week after birth and continues to destroy the infant's cells, though at a diminishing rate, all this time.

The best treatment for severe hæmolytic disease is an *exchange transfusion* carried out soon after birth. A catheter is passed along the umbilical vein into the inferior vena cava; small quantities of the infant's blood are successively withdrawn and replaced by an equal volume of compatible *Rh negative blood.* The infant's Rh+ red cells which were doomed to destruction are thus removed from the circulation and the infant's organs do not have to deal with the products of their disintegration; the infant is given an adequate

supply of Rh— red cells which will survive in the circulation for the normal length of time.

Rh Factor and Blood Transfusion.—The following rules should be observed.

(1) No Rh— female at any age before the menopause should ever be given a Rh+ blood transfusion if it can possibly be avoided. If she is Rh— she becomes sensitized by the injected Rh+ blood and forms anti-D; she is likely to destroy, subsequently, any Rh+ fœtus with which she becomes pregnant. In other words the transfusion may make her permanently sterile.

(2) Any Rh— woman who has given birth to a child suffering from hæmolytic disease is undoubtedly immunized to D and her serum contains anti-D. Likewise any Rh— person who is given Rh+ blood may become sensitized and the serum may contain anti-D. Anyone whose serum contains anti-D will show all the changes attributable to mismatched transfusion if they are transfused with a blood, *otherwise compatible, but containing D.*

(3) To avoid sensitization it is desirable that Rh as well as ABO testing should be carried out whenever possible before a blood transfusion is performed, whatever the sex of the recipient.

DIRECT CROSS-MATCHING BEFORE TRANSFUSION.—The only sure safeguard against transfusion complications is to match the serum of the recipient directly against the cells of the donor. The donor's cells should be diluted with a 20% albumin solution; under these circumstances the so-called anti-Rh albumin agglutinin (p. 179) can also be demonstrated. If the serum does not agglutinate the donor's cells, the donor's cells are compatible; if the serum agglutinates the donor's cells, his blood is incompatible.

USE OF PLASMA FOR TRANSFUSION.—When comparatively small volumes of plasma are used its agglutinin content can be ignored because its titre is reduced to harmless levels by dilution in the recipient's plasma. If large volumes of plasma are employed the injected agglutinins may be harmful. It is then wiser to use plasma from which the agglutinins have been removed by contact with the corresponding red cells (*conditioned plasma*). If a blood bank is used into which blood of all four groups is placed, the α and β agglutinins are absorbed by the A, B, and AB cells and thus removed.

Effects of Incompatible (Mismatched) Blood Transfusions.—When whole blood is injected intravenously into patients (*blood transfusion*), serious symptoms and even death can occur if the recipient's serum contains antibodies (α, β, or anti-D) which agglutinate the donor's red cells. The red cells are first *agglutinated* and then undergo *hæmolysis.* The following types of clinical reaction occur:

(i) INAPPARENT HÆMOLYSIS.—The injected red cells are rapidly destroyed, the recipient's blood returning within a week or less to its pre-transfusion state. No other symptoms are observed.

(ii) POST-TRANSFUSION JAUNDICE.—The injected red cells are destroyed and the bilirubin formed from the released hæmoglobin accumulates in the blood in sufficient amounts to produce jaundice (*hæmolytic jaundice*). The amount of urobilinogen in the urine is increased.

(iii) SEVERE REACTION WITH HÆMOGLOBINURIA AND RENAL FAILURE.[1]—Soon after beginning the transfusion, perhaps after a few c.c. of blood have

[1] Vavasquez, *J. Path. Bact.*, 1940, *41*, 413.

been introduced, the patient complains of violent pain in the back or elsewhere, and tightness in the chest; these symptoms are attributed to the agglutinated red cells forming clumps which block capillaries. The masses are then hæmolyzed; the released hæmoglobin colours the plasma bright red; some of the hæmoglobin is converted into the brown pigment methæmalbumin (p. 187). With rapid destruction of the hæmoglobin, bilirubin accumulates in the plasma and stains the tissues (*jaundice*).

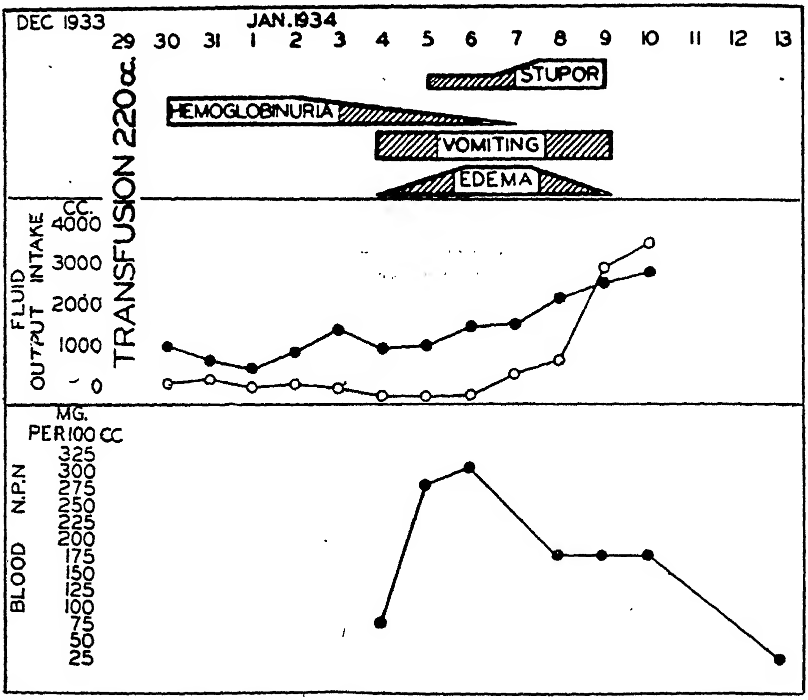


FIG. 101.—Results of Incompatible Blood Transfusion.

Records from above downwards: clinical signs; fluid intake (black dots), urine output (clear circles); blood non-protein nitrogen (N.P.N.) in mg. per 100 c.c.
The patient (recipient) belonged to Group O (serum=$\alpha\beta$); the donor's cells belonged to group A (Goldring and Graef, *Arch. int. Med.*, 1936, *58*, 838.)

The volume of urine is greatly decreased. This is due in part to a fall of arterial blood pressure (*e.g.* down to 70 mm. Hg) and perhaps to local vascular disturbances in the kidneys[1] such as closure of the interlobular arteries; the arterial blood is then shunted via the juxta-medullary glomeruli into the medulla and so does not pass through the majority of the cortical glomeruli, leading to grave reduction of the glomerular filtrate volume (p. 26). If the urine is acid and glomerular filtration is slow, the hæmoglobin which is passed out in the glomeruli is precipitated in the tubules possibly as acid hæmatin, owing to the filtrate becoming acid as it passes down the tubule; the lumen of the tubules is thus obstructed, and the flow of urine is

[1] Miller and McDonald, *J. clin. Investig.*, 1951, *30*, 1033.

seriously interfered with. In spite of alleviative measures such as injections of saline and alkali,[1] almost complete *anuria* may develop (Fig. 101). Owing to renal failure the nitrogenous constituents of the blood rise, and a condition of *latent uræmia* results (p. 73). The patient in severe cases may die in 8–10 days, after developing increasing lethargy and coma.[2] In a series of cases in which incompatible blood of the ABO groups was given, no death occurred when less than 350 c.c. of blood were transfused.

Inheritance of Classical Blood Groups.—The four classical blood groups depend on three genes named after the corresponding factor, A, B, and O. Each person's blood group is determined by the two genes which he receives, one from each parent. Genes A and B are of equal strength, and when present together work independently. Gene O, on the other hand, is recessive, *i.e.* when coupled with A or B it does not function. Accordingly:

If a person receives genes	A + A *or* A + O	B + B *or* B + O	A + B	O + O
His blood group is .	A	B	AB	O
His genotype is . .	AA or AO	BB or BO	AB	OO

This information can be used in investigating cases of disputed paternity. A baby must receive one of three possible genes (A, B, or O) from each parent. Further, each parent transfers one of two genes to the child: an A parent (genotype AA or AO) can give A or O, a B parent (genotype BB or BO) B or O, an AB parent A or B, and an O parent (OO) O only.[3]

It follows that:

If the baby's group is—	Parents must have given it—	So if mother was—	Father could *not have been*—
O	O + O	No matter which	AB
AB	A + B	No matter which	O
A	A + O or A + A	B or O	B or O
B	B + O or B + B	A or O	A or O

The medico-legal value of blood grouping tests is greatly enhanced if the MN and the various Rhesus factors are also studied.

The M and N factors depend on two minor supplementary blood genes.

[1] Owing to failure of salt and water excretion there is a real danger of overtreatment with the risk of fluid retention and œdema, especially of the lungs. Attempts at "flushing out the kidneys" should not be made; the volume of fluid administered should equal the maximum likely to be lost from the skin and lungs, and excreted by the kidneys.

[2] Whitby, *Lancet*, 1942, ii, 581.

[3] It must be remembered however that the child's blood may not be set in its true ABO type until as late as one year after birth.

These curiously enough do not elicit an output of agglutinins when injected into man and may therefore be ignored in carrying out transfusions. They are, however, antigenic to rabbits and agglutinating sera can be prepared by injecting human M or N cells into rabbits.

Each person carries two of the genes of the M and N group, *i.e.* M+M (=M), N+N (=N), or M+N (=MN). As a result:

If a baby's supplementary group is—	Parents must have given it—	So if mother's supplementary group was—	Father could *not have been*—
M	M + M	No matter which	N
N	N + N	No matter which	M
MN	M + N	N	N
MN	M + N	M	M

It should always be remembered that blood grouping tests can never prove that any suspected person *is* the actual father; they can only show that he could not possibly have been the father or that he (like many others) might have been.

Use of Stored Blood.—*Whole blood* transfusion is the ideal treatment for severe hæmorrhage when given promptly and in adequate amounts. Under war conditions and even for the emergencies of civil practice it is advisable to have considerable stores ready at hand. Group O blood is employed; it is collected and stored under strictly sterile conditions and kept fluid by addition of citrate. Group O, Rh— blood should be available for use in appropriate cases. [The $\alpha\beta$ agglutinins can be removed from the blood by adding the AB specific substances.]

Red cells undergo rapid changes during storage in simple citrate solutions; they are preserved much longer in the presence of 1% glucose, the amount of hæmolysis after six weeks' storage being reduced to one-tenth.[1] Glucose acts mainly by liberating lactic acid and lowering the *p*H of the medium (citric acid is helpful for the same reason). A satisfactory diluent is: 100 c.c. of 2% disodium citrate and 20 c.c. of 15% glucose, to which are added 420 c.c. of blood.

The main red cell changes during storage are fully described on p. 7; owing to the decrease or arrest of red cell metabolism the ions to which the cell membrane is permeable move according to their concentration gradients: (i) K^+ passes out of the cells into the plasma where the K^+ concentration rises from the normal 20 mg–% to, *e.g.*, 116 mg–% in 14 days (Fig. 102); Na^+ passes from the plasma into the cells, raising the concentration there from 40 mg–% to 80 or even 140 mg–%; (ii) the cells increase in volume and become shorter and fatter, *i.e.* more spherocytic (p. 164): in consequence they undergo hæmolysis more readily in hypotonic solution and may rupture *in vitro* in salt concentrations as high as 0·8–0·85%; (iii) hæmolysis of the cells takes place to an increasing degree while in contact with their own plasma; (iv) the inorganic phosphate of the plasma rises slightly from 3 mg–% to say 5 mg–%. The changes in the other blood constituents are unimportant:

[1] Maizels, *Quart. J. exp. Physiol.*, 1943, *32*, 143.

the platelets disappear in two days; 70% of the white blood cells have disappeared on the second day and many of the surviving ones are degenerate; by 10 days all the white cells are gone or dying; the constituents related to blood clotting and the plasma proteins are comparatively stable.

From the practical point of view the value of stored blood is closely related to the survival time *in the recipient's body* of the transfused cells. If they disintegrate very rapidly they create the difficulties always associated with free hæmoglobin in the plasma (p. 182) and do not help in oxygen carriage; such a whole blood is no better, and in obvious respects more dangerous, than plasma. The degree of spherocytosis and saline fragility of the cells prior to injection is *not* a reliable criterion of the fate of the cells in the body. Their survival time in the body must be determined by direct study of the recipient (Fig. 103). It is found that cells stored in citrate solution for 7 days disappear from the recipient's circulation in 1–2 days; by contrast, cells stored in citrate-glucose for as long as 18 days undergo no lasting deterioration apart from normal ageing, and they survive in the recipient's circulation (allowing for ageing) for as long as fresh cells.

As explained on p. 8, if abnormal stored red cells are incubated with glucose *in vitro* at 37° C. their metabolism is restored; consequently the Na^+ is extruded from the cells, the K^+ is drawn back and the ionic pattern of the red cells returns to normal. Similarly, if abnormal stored red cells are injected into the circulation of a recipient they become normal ("reconditioned") in less than 24 hours with respect to Na^+ and K^+ content, volume, shape and saline fragility.

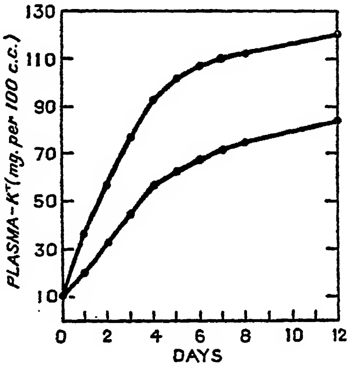


FIG. 102.—Changes in Plasma-Potassium Concentration in Stored Blood on Standing (temperature, 2–4° C.).

Ordinate: plasma-K^+ in mg./100 c.c. Note progressive increase in plasma-K^+ with time. The results with two bottles of blood are illustrated. (Downman *et al.*, *Brit. med. J.*, 1940, i, 559.)

The high plasma-K^+ concentration of stored blood must be borne in mind, as K^+ in excess has toxic effects on the circulation; these may be avoided if the transfusion is performed at rates not exceeding 40 c.c. per minute.

Plasma can be stored in the *liquid form for many months* [1]: if dried it can be kept for years and under all conditions of temperature; it can be reconstituted by adding sterile water. Plasma is of great value in hæmorrhage (though not quite as satisfactory as whole blood (p. 86)), traumatic shock, and burns. Conditioned plasma (p. 181) can be given in very large volumes without fear of agglutinating the recipient's red cells. The plasma is prepared from large pools representing many donors and some batches may carry the

[1] Stored plasma may occasionally develop toxic properties owing to some alteration taking place in its proteins. On intravenous injection it then produces a slowing of the heart and a fall of blood pressure (due to cardiac inhibition and vasodilatation) and disturbances of respiration. The toxic plasma seems to damage the lungs, stimulates local vagal endings and sets up afferent impulses which reflexly produce the circulatory and respiratory changes mentioned; the changes are prevented by section of the afferent pulmonary vagal fibres. (Gilding and Nutt, *J. Physiol.*, 1944, *102*, 446.)

virus of infective hepatitis. For this reason transfusion with the polysaccharide "Dextran" (6% in saline), is preferred by many. The large molecules of this substance are well retained in the circulation and help to maintain the blood pressure.

FATE OF THE RED BLOOD CORPUSCLES. JAUNDICE.

Duration of Life of Red Cells.[1]—Once the reticulocyte stage has been passed there are no means of determining how old any particular erythrocyte is. The average duration of life of erythrocytes in man is 3 or 4 months (and

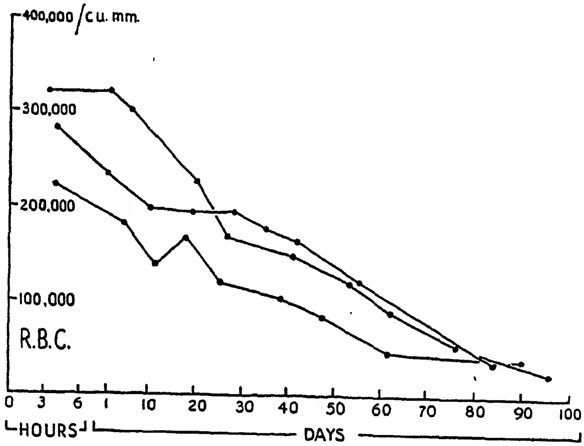


FIG. 103.—Survival Time of Transfused Red Blood Cells in Man.

Blood of Group O was transfused into patients of Groups A, B, or AB. Red-cell counts were performed after agglutinating the recipient's own cells with $\alpha\beta$ sera. The ordinate represents counts of Group O cells only in recipient's blood. Note that the injected cells may survive for 100 days. (Scarborough, *Edin. med. J.*)

not 3 or 4 weeks as previously supposed). The evidence is as follows: a large volume of Group O blood is transfused after a hæmorrhage into (say) a Group AB recipient. Counts of Group O cells are performed at intervals by diluting specimens of the recipient's blood with fluid containing serum α and β which agglutinate and precipitate the AB cells which are then filtered off. Typical results are shown in Fig. 103; the transfused cells progressively decrease in number but some are still present after 70–100 days.

Red Cell Survival Time in Disease.[2]—The survival time of red cells in the body is reduced when (i) the cells are abnormal or (ii) the hæmolytic system is abnormal (*i.e.* the normal destroying mechanisms are excessively active or abnormal hæmolysins are present).

(i) (*a*) The red cells of patients with congenital hæmolytic jaundice (p. 229) and pernicious anæmia (p. 196) are destroyed excessively rapidly in the

[1] London *et al.*, *J. biol. Chem.*, 1949, *179*, 463.

[2] Mollison, *Clin. Sci.*, 1946–48, *6*, 137. Singer and Motulsky, *J. lab. clin. Med.*, 1949, *34*, 768.

patient's body, thus producing or contributing to the anæmia. If these cells are transfused into normal subjects they are also rapidly destroyed. If normal cells are transfused into the patients their survival time is normal. These observations prove that in congenital hæmolytic jaundice and pernicious anæmia the red cells are abnormal but the hæmolytic mechanisms are normal.

(*b*) Red cells that have been stored *in vitro*, especially in unsuitable diluting fluids (p. 184), become abnormal and their survival time after transfusion is shortened.

(ii) In other forms of hæmolytic jaundice the red cells are normal but are rapidly destroyed by abnormal hæmolytic mechanisms. Normal red cells transfused into these patients have a short survival time; the cells of the patients when transfused into normal subjects have a normal survival time.

Destruction of Red Cells.—The initial stages in the destruction of the red cells are uncertain. The following suggestions have some evidence in their support:

(i) Intact red cells (presumably the aged ones) are taken up intact by the *macrophages* (cf. Fig. 126): this process seems to be uncommon in normal man.

(ii) The cells as they age are "threshed to bits" in the circulating blood and the fragments (fine hæmoglobin-containing "dust") are ingested by the macrophages.

(iii) The red cells may undergo hæmolysis in the blood stream, possibly as a result of preliminary treatment in the spleen (p. 227); the released hæmoglobin may (*a*) be taken up as such by the macrophages, or (*b*) break down in the blood stream into hæm and globin; the hæm then unites with serum albumin to form *methæmalbumin* which is ingested by the macrophages.

(iv) The hæmoglobin may undergo chemical changes while still within the red cells; thus the hæm may lose its iron, a bilirubin-globin complex resulting (called *verdohæmoglobin*) which may escape from the cell and be taken up by the macrophages.

Ultimately, whatever the intermediate stages may have been, the red cell is destroyed and the hæmoglobin is released; it is broken down to globin and an iron-free pigment called *bilirubin*. As is explained below bilirubin formation occurs mainly in the macrophage (reticulo-endothelial) system. The released iron is most carefully stored in the body probably by becoming bound with a special tissue protein called *apoferritin* to form the iron-containing protein, *ferritin* (p. 204). This "reserve iron" is presumably released into the circulation for hæmoglobin formation in the red marrow as and when required.

Macrophage [Reticulo-Endothelial] System.—The term reticulo-endothelial system is used in this book to refer to certain phagocytic cells found mainly in the bone marrow, liver, lymph nodes, spleen (and subcutaneous tissues). In the marrow the cells form part of the lining of the blood sinuses (*littoral* cells); in the liver they lie at intervals along the vascular capillaries (*Kupffer* cells); in the lymph nodes they line the lymphatic paths; in the spleen they are found in the pulp (p. 223). The characteristic feature of the cells of the reticulo-endothelial system is their power to ingest foreign colloidal particles; thus if carmine or Indian ink is injected intravenously into living animals the cells take up the dye, are deeply stained

by it and so are easily recognized. Cells of similar histological appearance which do not take up these dyes are not included in the reticulo-endothelial system. Because they ingest large particles the cells are called *macrophages*. The functions of these cells are as follows:

(i) They ingest and destroy red blood corpuscles and form and release bilirubin (*infra*).

(ii) They ingest bacteria and are thus concerned with the defence of the body against infection. They rapidly increase in number under these conditions with resulting enlargement of the organs which are rich in these cells, *e.g.* spleen, lymph nodes.

(iii) They form antibodies in response to the presence of certain types of protein antigens; it is suggested that fragments of the superficial cytoplasm are released containing the specific globulin antibody.

The reticulo-endothelial system functions as a physiological unity; if any part of it is put out of action the rest of the system undergoes compensatory hypertrophy and makes good the deficiency.

Origin of Bilirubin.—(1) In the normal animal, and particularly when blood destruction is actively proceeding, the macrophages (reticulo-endothelial system) contain fragments of red cells (Fig. 126), free hæmoglobin, or iron in an inorganic form which gives the Prussian blue reaction.

(2) There is a trace of bilirubin in the circulating blood; but the blood leaving the *spleen* and the *bone marrow* contains significantly higher concentrations of bilirubin than the arterial blood, proving that bilirubin is formed in these organs.

(3) Although the Kupffer cells of the liver also make bilirubin, the liver is *not* indispensable for bilirubin formation. After total extirpation of the liver, bilirubin rapidly accumulates in the body; the plasma and body fat are coloured yellow, and jaundice is present in animals which survive for more than six hours (cf. p. 825). Jaundice develops almost as rapidly when the spleen and the other abdominal viscera are removed as well, proving that the bone marrow macrophages are important sites of bilirubin formation.

(4) After *splenectomy* (in the intact animal), compensatory hypertrophy of the other macrophages takes place, and the rate of bilirubin formation is not diminished.[1]

Chemistry of Bilirubin Formation.—The first change undergone by hæm (or to give it its full chemical name, iron-protoporphyrin IX) is the oxidation of the C of its α methine ($\equiv$CH) bridge, to CO_2. The tetrapyrrol ring structure is thus broken up; the four pyrrol groups then become arranged as a *straight chain*. Next, the Fe is split off, the resulting compound being bilirubin. The sequence of events is set out in Fig. 104, which also shows that bilirubin has the same side chains attached to its pyrrol groups as has hæm; it differs from hæm in the loss of the $\equiv$CH (methine) bridge in the α position and in being a straight chain and not a ring structure.

The bilirubin which is released by the macrophages into the blood stream may be called *hæmobilirubin*; in the circulation it is probably carried in combination with serum protein. It is secreted in a modified form (probably

[1] In *bruises*, after the passage of time the hæmoglobin of the extravasated blood is converted progressively into bilirubin owing to the activity of the local macrophages; these, however, are not concerned with the normal destruction of circulating blood cells.

split off from protein) by the hepatic cells into the bile passages; the bilirubin in the bile is called *cholebilirubin.*

VAN DEN BERGH REACTIONS.—These reactions are of value because they enable the two forms of bilirubin, *i.e.* hæmobilirubin and cholebilirubin, to be distinguished. They are based on the fact that when a mixture of

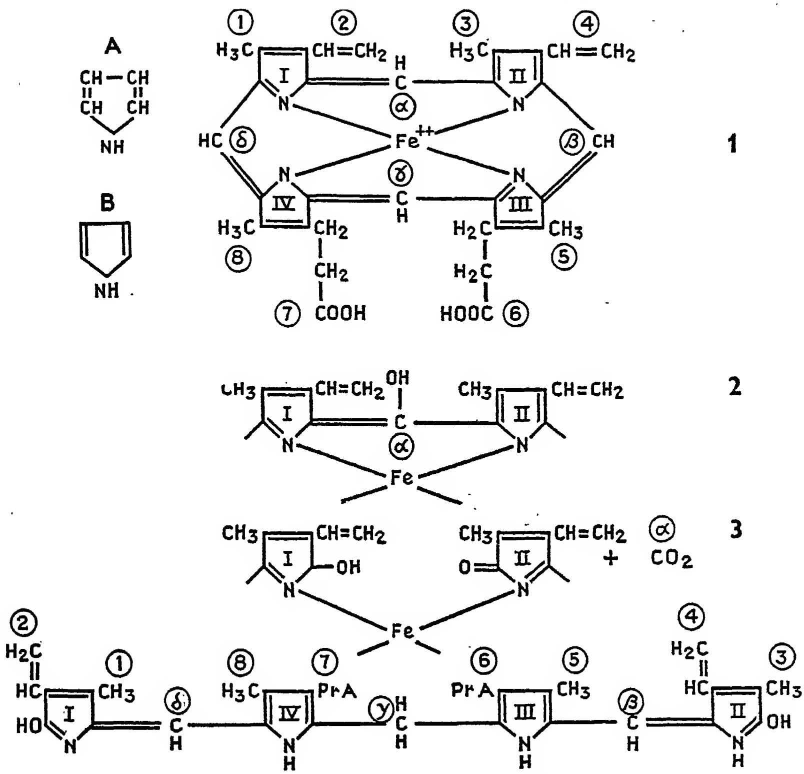


FIG. 104.—Chemistry of Bilirubin Formation. (After Granick, *Ann. N.Y. Acad. Sci.*, 1946–47, *48*, 657.)

A. Pyrrol ring in full.
B. Pyrrol ring, conventional outline.
1. Hæm (iron-protoporphyrin IX). In 2 and 3 only rings I, II, and αC are shown. 2. Oxidation at αC of CH to COH. 3. Further oxidation and loss of αC. The Fe atom is subsequently split off. Lowest formula is that of bilirubin.

sulphanilic acid, HCl, and sodium nitrite is added to bilirubin, a reddish or reddish-violet compound (*azobilirubin*) is formed. Two kinds of reaction can be obtained: the *direct* and the *indirect.*

(1) *Direct Reaction.*—The serum is treated directly with the reagents. A positive response consists of the appearance of the characteristic colour: (i) *rapidly, i.e.* the maximum colour intensity is attained within 30 seconds; this is the *prompt* direct reaction; or (ii) the colour develops slowly but is obvious within 10 minutes; this is the *delayed* direct reaction. If no colour

appears within 10 minutes the reaction is negative. A positive direct reaction is given by cholebilirubin, *i.e.* the bilirubin which has been secreted into and is present in bile, or the bilirubin which has passed back from the bile into the blood. Hæmobilirubin gives a negative direct Van den Bergh reaction.

(2) *Indirect Reaction.*—The serum and the reagents are mixed and then *alcohol* is added which precipitates the serum proteins. A reddish colour indicates the presence of bilirubin. *Both* forms of bilirubin give a positive reaction.

The *qualitative* Van den Bergh reactions are less helpful clinically than the corresponding *quantitative* reactions. When the latter are used the indirect reaction determines the *total* amount of *both* forms of bilirubin; the direct reaction gives the amount of cholebilirubin; the difference between the two values represents the amount of hæmobilirubin.

The main differences between the two kinds of bilirubin are summarized below:

HÆMOBILIRUBIN.	CHOLEBILIRUBIN.
Normally present in blood-serum.	Normally present in bile.
Increased in amount in hæmolytic jaundice ("retention jaundice").	Appears in serum in obstructive jaundice ("regurgitation jaundice").
Gives a negative direct Van den Bergh reaction.	Gives a positive direct Van den Bergh reaction.
Gives a positive indirect Van den Bergh reaction.	Gives a positive indirect Van den Bergh reaction.
Extractable by chloroform.	Not extractable by chloroform.
Present in serum in the form of (?) bilirubin-albumin.	Present in bile in the form of (?) Na bilirubinate.
Not excreted in the urine.	Excreted in the urine.

SERUM BILIRUBIN VALUES.—The normal adult range is 0·2–1·7 mg–%; (in 93% of normal cases the level is below 0·8 mg–%); values over 1 mg–% are unusual; the mean is 0·5 mg–%. In new-born infants higher values are obtained (p. 192).

Jaundice.—In jaundice there is an abnormally raised bilirubin content of the blood, with resulting outflow of bilirubin into the tissues, which consequently become stained yellow. Three main varieties of jaundice may be recognized:

(1) HÆMOLYTIC JAUNDICE.[1]—This results from excessive destruction of red blood cells, with consequent formation of bilirubin more rapidly than the hepatic cells can cope with it; excess hæmobilirubin thus accumulates in the blood ("retention jaundice"). Rapid hæmolysis may occur after transfusion of incompatible blood (p. 181), experimental injection of hæmolytic sera (p. 170) or of poisonous agents like arseniuretted hydrogen, in pernicious anæmia (p. 196), or in congenital hæmolytic (acholuric) jaundice (p. 229).

In hæmolytic jaundice, although bilirubin accumulates in the blood, the bile acids are passed out normally into the intestine and their concentration in the blood does not rise. Hæmolytic jaundice is therefore a "pure pigmentary cholæmia" or "*dissociated jaundice*"—the bile salts being treated differently from the pigments.

[1] Davis, *Edin. med. J.*, 1943, *50*, 589.

Other factors may contribute to the production of jaundice in these cases: (i) The hæmolytic agent may damage the liver cells, impairing their power to secrete bilirubin from the blood into the bile; the liver cells may become swollen, blocking the bile canaliculi. (ii) The raised bilirubin content makes the bile very viscid and may lead to the formation of bile thrombi; the flow of bile is slowed down, with resulting rise of pressure in the bile passages and back-flow of some cholebilirubin into the blood. In uncomplicated hæmolytic jaundice, bile pigments are absent from the urine, but an excessive amount of urobilinogen (p. 194) is excreted.

(2) Toxic and Infective Jaundice.—The liver cells are poisoned and consequently their power of secreting bilirubin from the blood into the bile

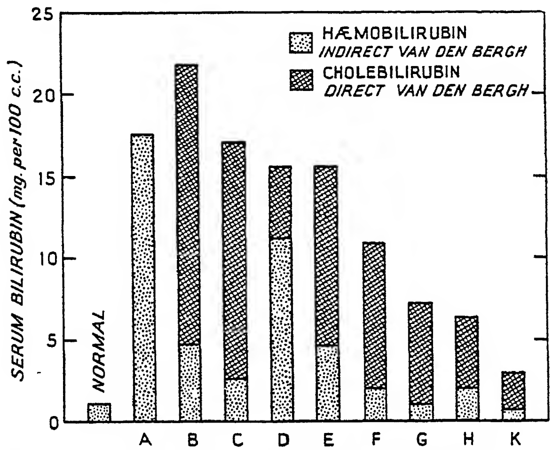


Fig. 105.—Distribution of the two types of Bilirubin in the Serum in an Illustrative Series of cases of Jaundice. (Modified from Green *et al.*, *Arch. int. Med.*, 1938, *61*, 658.)

A. Hæmolytic Jaundice; B. Hepatitis; C. Hepatitis and Cirrhosis; D. Stricture of Bile Duct; E, F. Stone in Common Bile Duct; G. Stone in Bile Duct and Cirrhosis; H. Carcinoma of Liver K. Portal Cirrhosis.

capillaries is impaired; hæmobilirubin is consequently retained in the blood.

Frequently an obstructive factor is also present: (i) because the liver cells are swollen they tend to block the bile canaliculi; (ii) the inflammatory process may involve the bile ducts; the bile capillaries may thus become plugged in places by an albuminous coagulable material. The distended bile capillaries above the block may even rupture with resulting return of cholebilirubin into the blood.

(3) Obstructive Jaundice.—In this condition there is an obstruction to the outflow of bile from the bile passages into the small intestine. Bilirubin which has already been secreted from the blood into the bile, *i.e. cholebilirubin*. is passed back into the blood where it accumulates ("regurgitation jaundice"). Obstructive jaundice may be due to some gross cause such as a stone in the common bile duct or a carcinoma of the head of the pancreas. In such

conditions *all* the bile constituents (pigments and salts) are returned into the blood and do not reach the intestine. Bile pigments appear in the urine when the serum level exceeds 2 mg–%.

As explained above there may be an obstructive factor in hæmolytic or toxic and infective jaundice. In some grave affections of the liver such as acute hepatitis (p. 833), widespread necrosis of liver cells occurs. There is then not only secretory failure but many liver cells become cut off from the bile capillaries and any bilirubin that is still secreted is passed back as cholebilirubin into the blood.

From what has been said above it is clear that, although cases of jaundice can be usefully divided into three main groups, considerable overlapping occurs, more especially between the second and third groups. Thus, in "pure" toxic and infective jaundice, hæmobilirubin alone should accumulate in the blood, giving a positive indirect reaction only; in "pure" obstructive jaundice, only cholebilirubin should accumulate in the blood, giving both a positive direct as well as an indirect reaction. Fig. 105, however, presents the actual blood findings in some illustrative cases of jaundice. In the case of hæmolytic jaundice studied hæmobilirubin alone was markedly increased, as would be expected (Fig. 105, A). In the cases of liver damage (hepatitis, carcinoma, cirrhosis (Fig. 105, B, C, H, K), *cholebilirubin* as well as hæmobilirubin was present in excess, and unexpectedly the blood generally contained more cholebilirubin than bilirubin. In the cases of biliary obstruction the hæmobilirubin as well as the cholebilirubin was raised (Fig. 105, D, E, F; in G the hæmobilirubin was normal); in fact in the examples shown there was surprisingly little difference in the blood bilirubin picture between the cases of liver damage and of biliary obstruction. (For diagnostic value of *flocculation tests*, see p. 831.)

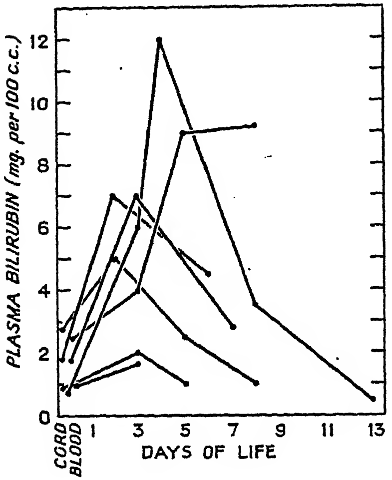


FIG. 106.—Plasma Bilirubin Content in Blood of Umbilical Cord and Post-Natal Life. (Findlay, Higgins, and Stanier, *Arch. Dis. Childh.*, 1947, *22*, 65.)

A series of normal infants was examined. When the plasma bilirubin exceeds 5 mg–% clinical jaundice is present.

JAUNDICE (HYPERBILIRUBINÆMIA) OF THE NEW-BORN (ICTERUS NEONATORUM).[1]—As mentioned on p. 190, it is unusual for the plasma bilirubin concentration in normal adults to exceed 1 mg–%; higher values are, however, very common in normal young infants. In 68 of 110 normal new-born infants the plasma bilirubin in the umbilical cord blood exceeded 1 mg–%. In many infants the bilirubin level continues to rise after birth to a peak which is generally reached during the first week, and then declines; if the plasma bilirubin exceeds 5 mg–%, clinical jaundice is always present. In 73 normal infants examined, the peak bilirubin level attained during the

[1] Findlay, Higgins, and Stanier, *Arch. Dis. Childh.*, 1947, *22*, 65.

first 10 days exceeded 1 mg–% in 34; it exceeded 2 mg–% in 28 of these, the highest value noted being 20 mg–%; 18 cases showed clinical jaundice, the so-called *icterus neonatorum* (Fig. 106). The cause of this "physiological" jaundice of the new-born has been much discussed.

(1) The jaundice has been attributed to excessive hæmolysis. In support of this view it is emphasized that a rapid fall of red cell and hæmoglobin concentration normally occurs after birth; the fall is maximal during the second week and may continue into the third month; as a result the blood picture changes from one of polycythæmia (by adult standards) to one of anæmia (Fig. 100). But the usual evidence of excessive hæmolysis is absent; thus no hæmolysins have been demonstrated and there is no raised saline fragility of the red cells; there is no relationship between the time of onset of the fall or the rate of fall of the red cell count and the onset or the severity of the jaundice. The fall in the red cell count immediately after birth is probably due to *decreased activity of the red marrow* and not to increased hæmolysis.

(2) The jaundice of the new-born is probably due to *hepatic immaturity*. Many physiological functions are imperfectly carried out in the new-born, *e.g.* temperature regulation, voluntary movements. *In utero* the bilirubin formed is mainly eliminated via the *placenta*; it is secreted to a lesser extent by the liver into the bile and reaches the intestine to form the green *meconium*. Immediately after birth the liver has to eliminate *all* the bilirubin formed; it would seem that frequently the liver is unable to deal adequately with this task during the first 10 days of life; jaundice therefore develops. Bilirubin excretion in the fæces in these jaundiced infants is always decreased.

EXCRETION OF BILE PIGMENTS IN URINE.—Hæmobilirubin is not excreted by the kidney, but cholebilirubin is excreted when the plasma level exceeds 2 mg–%. In general then, hæmolytic jaundice tends to be *acholuric* (p. 229), *i.e.* there is no bile pigment in the urine; in obstructive jaundice, however, bile pigment (and bile salts) appear in the urine in amounts which are proportional to the concentration of these substances in the blood.

Fate of Bile Pigments—In the bile passages bilirubin is oxidized by the loss of 2H to *biliverdin*: the pyrrol side chains are unaffected. Reduction of the bile pigments occurs in the intestine to *urobilinogen* (*stercobilinogen*).

IN INTESTINE IN AIR

$$\begin{array}{lccc}
\text{Biliverdin} & & & \\
\uparrow -2\text{H} & & & \\
\text{Bilirubin} \xrightarrow[+12\text{H}]{} & \begin{array}{c}\text{Urobilinogen}\\ \text{(Stercobilinogen)}\end{array} & \xrightarrow[-2\text{H}]{} & \begin{array}{c}\text{Urobilin}\\ \text{(Stercobilin)}\end{array}
\end{array}$$

The reduced pigment, urobilinogen (stercobilinogen), is excreted in the fæces; on exposure to air it is oxidized into urobilin (stercobilin). About half the pigment excreted in the bile can be demonstrated in the fæces. Some is reabsorbed from the intestine into the portal blood and reaches the liver where it is largely re-excreted into the bile, although a little escapes into the systemic circulation to be excreted in the urine as urobilinogen (which on exposure to air becomes urobilin). It is alleged that liver damage allows more of the absorbed urobilin to pass through; the amount of urobilinogen in the urine is thus increased. The amounts of bile pigment

derivatives normally excreted in the fæces is 40–280 mg., compared with 0–4 mg. in the urine.

In hæmolytic anæmia, because of excessive formation of hæmobilirubin, there is increased excretion of cholebilirubin, and more urobilin (stercobilin) is formed in the intestine. As larger amounts of urobilin are reabsorbed into the blood there is an excessive excretion of urobilinogen in the urine.

It is clear that no particular care is taken in the body to preserve the pyrrols derived from the breakdown of hæm. This may be related to the fact that tetrapyrrols can be readily *synthesized* in the body from simple constituents like glycine (p. 880) or "acetic acid units" (p. 873).

REGULATION OF ERYTHROPOIESIS. PERNICIOUS ANÆMIA AND OTHER MACROCYTIC ANÆMIAS [1]

Regulation of Erythropoiesis.—The main factors influencing erythropoiesis can be classified as follows:

A. General Stimulants.—(1) *Anoxia.*—Oxygen lack, when sufficiently severe, stimulates red cell formation in the marrow and increases the red cell count. Anoxia is responsible for the polycythæmia at high altitudes (p. 445) or of congenital heart disease (p. 453), for restoration of the red cells after hæmorrhage (p. 84), and for the marrow hypertrophy which occurs in many anæmias. Conversely a *high oxygen pressure depresses erythropoiesis* (Fig. 107 and legend).

(2) *Thyroid.*—In thyroid deficiency (myxœdema), owing to depressed marrow activity, anæmia (macrocytic and hypochromic) commonly occurs which responds to thyroid medication (but not to iron or liver therapy) (Fig. 108). Thyroxine probably has no specific stimulating action on the red marrow but acts as a general metabolic stimulant, the marrow participating n the widespread tissue response.

(3) *Vitamin-C.*—Although many cases of clinical scurvy are anæmic it has been shown that pure *complete vitamin-C* lack does *not* give rise to anæmia.[2]

B. Factors Affecting Hæmoglobin Formation.—Certain raw materials must be available in adequate amounts, namely, *iron* for hæm (pp. 205 *et seq.*) and *protein* for globin formation (p. 210); *copper* is also needed in minute amounts (p. 213) (though its mode of action is obscure). The *porphyrin* of hæm can be readily synthesized in the body.

C. Maturation Factors.—(1) *Endothelial Cell→Proerythroblast* (*Megaloblast*).—Nothing is known about the factors concerned in this initial phase of red cell formation. Arrested development at the stage of the endothelial cell leads to *aplastic anæmia*. The areas where red marrow is normally found are fatty and devoid of erythroid cells; there is a progressive decline in the circulating total red cell volume (Fig. 49) and the red cell count; death ultimately results from anoxia.

(2) *Maturation of Nucleated Red Cells* (*Erythroblasts*).—For this stage of

[1] Whitby and Britton, *Disorders of the Blood*, London, 6th edn., 1950 (full bibliog.).

[2] Animal experiments suggest that certain members of the vitamin-*B* group are involved in red cell formation, *e.g.* pyridoxine, riboflavin, and nicotinic acid. There is no evidence that their absence is related to human anæmia.

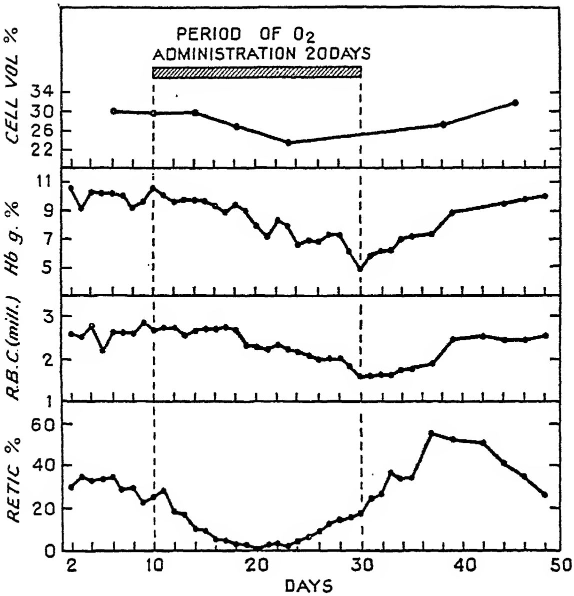


FIG. 107.—Effect of High Oxygen Concentrations on Erythropoiesis. (Moore *et al.*, *J. clin. Investig.*, 1944, *23*.)

A case of sickle-cell anæmia. In this disease the abnormal red cells are rapidly destroyed; the bone marrow becomes hyperplastic and young red cells enter the circulation prematurely leading to an increased reticulocyte count.

Initially the red cell count was 2·5 million, the hæmoglobin concentration 11 g–%, hæmatocrit value (red cell volume %) 30%, reticulocytes 20–30%. Pure O_2 was breathed through a mask for 20 days. During this period the reticulocyte count fell, indicating depressed erythropoiesis. The red cell count and hæmoglobin concentration fell because new red cell formation was not keeping pace with red cell destruction. On returning to air, intense erythropoiesis was resumed as shown by the increase in the reticulocytes (to over 50%) and the rise in the red cell count and hæmoglobin concentration.

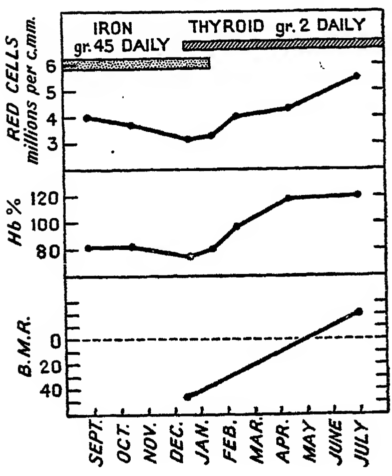


FIG. 108.—Action of Thyroid on Anæmia of Myxœdema. (Sharpe, *Amer. J. med. Sci.*, 1937, *194*, 385.)

Ordinates from above downwards: red cells in millions per c.mm.; hæmoglobin %; B.M.R. (basal metabolic rate): 0 = normal, 40 = 40% below normal.

Initial hæmoglobin 82%, red cell count 4·1 million per c.mm. On iron therapy (45 grains $FeNH_4$ citrate daily) the hæmoglobin fell to 76% and the red count to 3·2 million; mean corpuscular volume 92 μ^3, mean corpuscular hæmoglobin 41×10^{-12} g. After three months the B.M.R. was determined and found to be −37%. Iron therapy was discontinued and thyroid extract, 2 grains daily, given. The improvement in the blood roughly paralleled the rise of B.M.R. After six months the hæmoglobin was 120% and the red cell count 5·5 millions. The B.M.R. was 20% above normal.

development a specific maturating principle is necessary, the *hæmatinic principle*, which is formed in the stomach by the action of a gastric enzyme (*intrinsic* factor) on a food constituent (*extrinsic* factor). In the absence of hæmatinic principle the blood and marrow changes characteristic of *pernicious anæmia* (and other so-called *megaloblastic anæmias*) are found.

The evidence for the above statements is given on pp. 198 *et seq.*

Pernicious Anæmia.—This disease is due to the absence of the *hæmatinic* principle which regulates the maturation of the nucleated red cells in the marrow. The principal changes in pernicious anæmia are as follows:

(1) BONE MARROW.—Owing to the anæmia and the resulting anoxia the red marrow is stimulated, *i.e.* it becomes *hyperplastic* and spreads throughout the shafts of the long bones (*e.g.* femur, tibia, and fibula, and the corresponding bones of the arm). The nucleated red cells in the marrow are not only greatly increased in number but also show characteristic peculiarities. If the successive nucleated cell stages in erythropoiesis are called Stages I, II, III, and IV then in pernicious anæmia the *early* stages (I, II) *are more numerous* (*e.g.* 70%) than Stages III and IV (30%); normally the reverse is the case (p. 164). The nucleated red cells differ in their detailed cytology from normal marrow erythroid cells and are called by British hæmatologists, *megaloblasts*, early, intermediate, and late (p. 164). These cells are of *greater diameter*; the cytoplasm becomes *prematurely filled with hæmoglobin* and so becomes *eosinophil* sooner. The nucleus of the early, intermediate, and late megaloblast is more primitive than that of the corresponding normoblast and the chromatin is finer, sparser and less regularly spaced; much diagnostic significance is attached to these chromatin changes. The late megaloblast loses its nucleus and becomes a non-nucleated *large* red cell or *macrocyte*. Although the erythroid cells in the marrow of pernicious anæmia are very numerous most of them *fail to mature* beyond the early megaloblast stage; as the *complete* ripening of the nucleated cells is inhibited, fewer mature red cells enter the circulation, with resulting anæmia. The overactive marrow of pernicious anæmia is described as showing *megaloblastic hyperplasia.*

(2) BLOOD CHANGES.—The number of circulating red cells is greatly reduced, counts of 1 million cells per c.mm. not being rare. Fig. 94 shows that the average diameter of the cells is well above normal (*e.g.* 8·2 μ) and many cells may have diameters of 10 or 11 μ; the cells have a larger volume (95–160 μ^3 (average normal 87)), the mean corpuscular hæmoglobin concentration is the normal 35%, but the hæmoglobin content of each cell (mean corpuscular hæmoglobin) may be as much as 50×10^{-12} g. (normal 30). This means that the hæmoglobin concentration (in g-%) is reduced proportionately *less* than the red cell count (in millions per c.mm.). Nucleated red cells may sometimes be seen in blood films. The granulocytes and platelets are commonly reduced in number probably because of encroachment of the megaloblastic tissue on the space available in the marrow.

There is evidence that the abnormal circulating red cells are undergoing *excessive destruction*; the spleen, marrow, and lymph glands may show intense phagocytosis of red cells, the serum bilirubin may be elevated in consequence (the average concentration in pernicious anæmia is 0·98 mg. compared with 0·31 mg-% in normals) and a low grade of *jaundice* is present. If the cells of a patient with pernicious anæmia are injected into a normal person, their survival time is less than normal (p. 186.)

(3) Stomach.—There are regular and characteristic changes in the stomach, consisting of atrophy or destruction of that part of the mucosa which bears the peptic and oxyntic cells, but usually sparing the pyloric glands; complete *achlorhydria* is almost always present (pp. 202, 784).

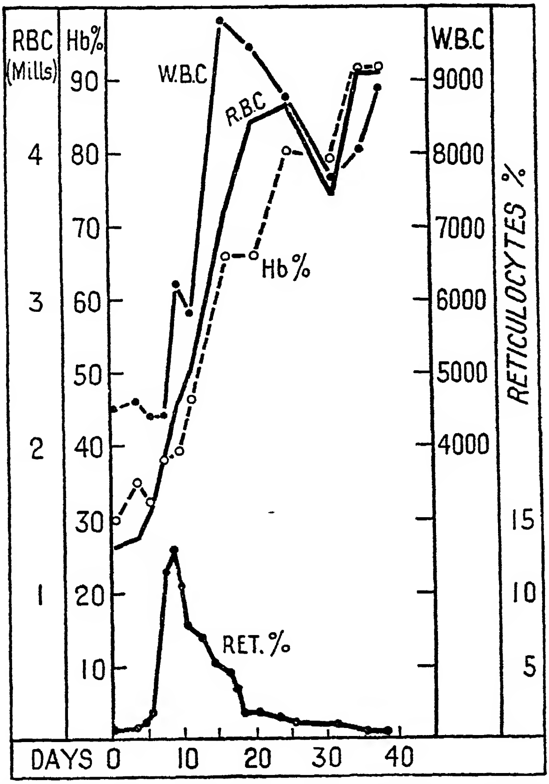


Fig. 109.—Liver Treatment of Pernicious Anæmia. (Brewer, Wells, and Fraser, *Brit. med. J.*, 1928.)

RBC: red cell count in millions per c.mm. WBC: white cell count in thousands per c.mm. Hb %: hæmoglobin % (Haldane scale). Ret. or Reticulocytes %: reticulocyte count per 100 red cells. ½-lb. cooked liver was given daily. Note the initial transient increase in the reticulocyte count which is followed by a progressive increase in the red cell count and hæmoglobin percentage. The white cell count also markedly rose.

(4) Nervous System.—In advanced cases demyelination of the white fibres of the spinal cord occurs, affecting the dorsal columns chiefly, and later the lateral columns (*subacute combined degeneration of the cord*).

Untreated cases show periods of remission when the blood picture improves followed by periods of relapse when the anæmia becomes more severe; but the general course is downhill until death occurs.

Cases suitably treated with active *liver extracts* given by mouth or by

intramuscular injection can be restored to, and maintained in, good health indefinitely. The active agent in the liver extracts is called the *hæmatinic principle.* A highly potent *cobalt*-containing crystalline substance (mol. wt. 1500) has been isolated from liver extracts; it is commonly called vitamin-B_{12}.[1] As little as 1 μg. given (*intravenously* or *intramuscularly*) daily restores the blood to normal and relieves the neurological symptoms.[2]

Action of Hæmatinic Principle in Pernicious Anæmia.—The first sign of improvement after commencing liver therapy is an increase in the number of circulating reticulocytes (*reticulocyte response*); within 7–12 days they reach a peak level and may constitute 15–40% of the total red cells. This response always heralds a general improvement in the condition of the circulating blood, and the hæmoglobin concentration and red cell count soon begin to rise. The number of reticulocytes then falls to normal, but the hæmoglobin and red cell levels continue to rise until the normal state of the blood is restored (Fig. 109).

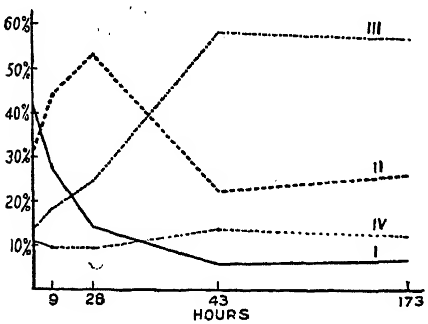


FIG. 110.—Effect of Liver Extract (Hæmatinic Principle) on Erythropoiesis in Pernicious Anæmia. (Davidson *et al.*, *Quart. J. Med.*, 1942, *11*.)

Ordinate: Percentage distribution of different types of nucleated red cells (I, II, III, and IV) in sternal *bone-marrow* films.
Abscissa: Hours after injecting liver extracts. Determinations were made at 9, 28, 43, and 173 hours (averages of 12 cases).
Stages I, II, III, and IV: successive stages in maturation of nucleated red cells (cf. p. 164). *Before* treatment, as the marrow is *megaloblastic*, stages I–IV include the early, intermediate, and late megaloblasts.
Stages I and II initially constituted 70% of the marrow erythroid cells: stages III and IV were 30%. *After* liver treatment stage I was less than 10% and stage III over 50%; the cells had also *reverted to the normoblastic type.*

MODE OF ACTION OF HÆMATINIC PRINCIPLE.[3]—(i) Within as short a period as 6–9 hours, and before there is any change in the circulating blood, the bone marrow begins to change its appearance (Fig. 110). The earlier megaloblastic cells (*i.e.* Stages I and II) which are initially the dominant cells become progressively transformed into the later stages (*i.e.* Stages III and IV) and all the marrow cells revert increasingly to the *normoblastic* type; in other words, hæmatinic principle promotes the *maturation* of the erythroid cells and enables them to *follow the normal line of development* as regards cell size and time of appearance of the hæmoglobin. As the red cells complete their maturation

[1] Given *by mouth,* 300–500 μg. are needed daily. For the chemistry of vitamin-B_{12} and detailed clinical reports see Lester Smith, *Brit. med. J.*, 1949, ii, 1367; Ungley, *ibid.*, 1370.

[2] *Folic Acid* (pteroyl glutamic acid), which is chemically quite unrelated to vitamin-B_{12} and is not intrinsic or extrinsic factor, restores the marrow and blood to normal in pernicious anæmia, but unfortunately *aggravates* or *induces the neurological changes.* It is, therefore, of no therapeutic value in this condition. Folic acid clinically restores the marrow and blood to normal in megaloblastic anæmias due to many causes: *e.g.* sprue, idiopathic steatorrhœa, pregnancy, and nutritional deficiencies; it is probably unwise to use it for other than short periods because of the danger of damage to the nervous system. The absence of folic acid from the diet of chickens produces a macrocytic hypochromic anæmia, leucopenia, and thrombopenia.

[3] Davidson, Davis, and Innes, *Quart. J. Med.*, 1942, *11*, 19.

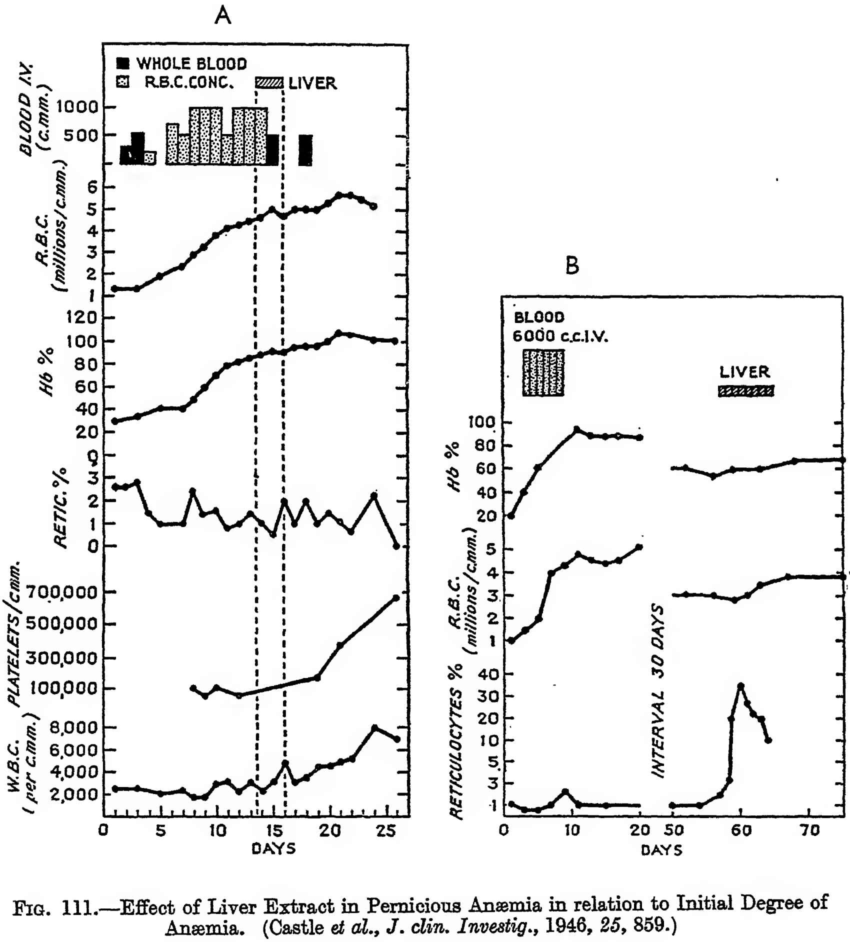


FIG. 111.—Effect of Liver Extract in Pernicious Anæmia in relation to Initial Degree of Anæmia. (Castle *et al.*, *J. clin. Investig.*, 1946, *25*, 859.)

A. Whole blood transfusion or concentrated red cells injected intravenously. Red cell count and hæmoglobin concentration were raised from 1·5 million and 30% to 4·5 million and 90% respectively. Inject liver extract as indicated. No reticulocytosis develops, though there is a further improvement in the red cell count. The platelet count and white cell count rise. R.B.C. concentrate is expressed as original volume of blood before removal of plasma.

B. Red cell concentrate corresponding to 6000 c c. of whole blood was injected intravenously. The red cell count and hæmoglobin concentration were raised from 1·0 million and 20% to 5·0 million and 90% respectively. About 30 days were allowed to elapse without treatment; the red cell count fell to 3·0 million and the hæmoglobin concentration to 60%. Liver extract was then injected. Note immediate reticulocyte response (from 1 to 35%) followed by slow recovery in red cell count and hæmoglobin concentration.

young reticulocytes begin to appear in the circulation (*reticulocyte response*) and subsequently the red cell count rises. The newly formed red cells are normal in size, *i.e.* they are normocytes and not macrocytes; they also no longer undergo excessively rapid destruction. The hæmatinic principle thus inhibits the process responsible for the undue peripheral hæmolysis possibly by promoting the formation of red cells of normal structure; the serum bilirubin correspondingly returns to normal.

(ii) Hæmatinic principle also acts on the *other marrow elements* as shown by the coincident increase in the white cell count (Fig. 109) and in the blood platelets.

(iii) Within a few hours of liver treatment the patient feels stronger, and some of the characteristic symptoms such as loss of appetite, apathy, and digestive disturbances disappear, though the red cell count is as yet unaltered and the anoxia is not relieved. If a transfusion of concentrated red cells is given to an untreated case of pernicious anæmia (to restore the red cell count to normal) the symptoms due to anæmia and anoxia (pallor, breathlessness, palpitation) disappear, but the sense of well-being is not restored. Hæmatinic principle thus seems to act not exclusively on the marrow but on cell metabolism in general.[1]

(iv) If hæmatinic principle is given to a patient in whom the circulating blood has been restored to normal by a transfusion, the typical marrow changes occur but there is *no reticulocyte response* (Fig. 111). In other words, hæmatinic principle *directly* promotes the maturation and normalization of the marrow cells, but the rapid *turning out* of red cells by the marrow depends on the stimulus of anoxia (which always induces marrow overactivity). Clinically it is found that the extent of the reticulocyte response is inversely related to the red cell count before treatment; if the initial count exceeds 3·5 million the reticulocyte response is slight.

(v) The changes in the stomach are unaffected by treatment.

(vi) As many nuclei of normoblasts are destroyed during the rapid recovery phase, there is increased excretion of uric acid in the urine from the disintegration of the nucleoprotein of these nuclei.

Formation of Hæmatinic Principle.—The mechanism of formation of hæmatinic principle can now be discussed. This involves consideration of the rôle of the changes in the gastric mucosa in pernicious anæmia. Castle carried out the following experiment: a normal person ate a mixed meal and allowed it to undergo preliminary digestion in his stomach; the digest was then aspirated and given by stomach tube to a patient with pernicious anæmia. If this "treatment" was continued a typical reticulocyte response and blood and clinical improvement resulted. The meal alone, or gastric juice alone, given to the patient had no curative action. Castle concluded that during digestion of a meal in a normal stomach the hæmatinic principle is formed by the interaction of a food constituent (*extrinsic* factor) and a gastric juice constituent (*intrinsic* factor). Thus:

Extrinsic Factor+Intrinsic Factor=Hæmatinic Principle.

Recent work, however, suggests that extrinsic factor is actually vitamin-B_{12} or some related food constituent and that intrinsic factor modifies it or interacts with it so as to *promote its absorption from the intestine.*

[1] Castle *et al.*, *J. clin. Investig.*, 1946, *25*, 858.

EXTRINSIC FACTOR.—Extrinsic factor is present in largest amounts in ox *muscle*, *yeast*, and yeast extracts like Marmite.[1]

INTRINSIC FACTOR.—A digest of ox muscle with whole normal gastric juice is curative, *i.e.* hæmatinic principle is formed. But if ox muscle is

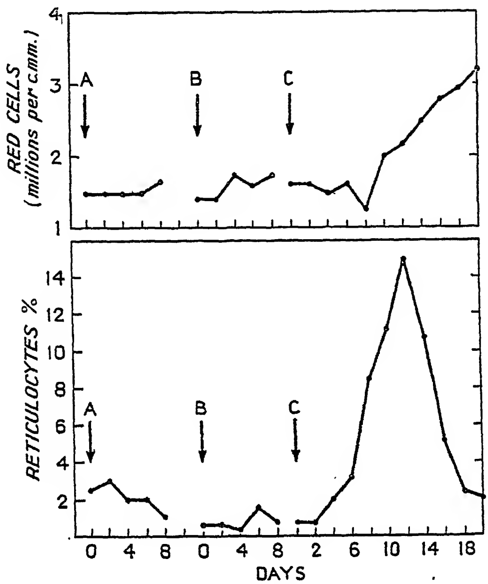


FIG. 112.—Experimental Analysis in a Case of Pernicious Anæmia of Formation of Hæmatinic Principle.

Upper records: red cell count (millions per c.mm.); lower records: reticulocytes %. The curative activity of the material employed is judged primarily by the *reticulocyte response* and secondarily and later by the rise in the red cell count.

A: administer by mouth 300 c.c. of normal gastric juice daily (=intrinsic factor): no response.

B: administer by mouth daily, in the morning, 200 g. of beef muscle (=extrinsic factor) incubated +HCl, and in the afternoon 300 c.c. of normal gastric juice (=intrinsic factor): no response. Thus intrinsic factor+extrinsic factor given separately at widely separated intervals are not curative.

C: administer by mouth daily 200 g. of beef muscle *previously incubated* with 300 c.c. of normal gastric juice: rapid full response.

Conclusion: Intrinsic Factor+Extrinsic Factor=Hæmatinic Principle.

(Drawn from data by Castle and Townsend, *Amer. J. med. Sci.*, 1929, *178*, 764.)

incubated with pure HCl no activity develops. Similar negative results are obtained on incubating muscle with other individual constituents of human gastric juice (*e.g.* pepsin, rennin, lipase), or with human saliva, or with pure human duodenal juice; and, of course, muscle given alone is inert. Intrinsic factor is thus a hitherto unrecognized constituent of human gastric juice and in its mode of action resembles an *enzyme* (Fig. 112).

[1] For recent studies see Hall, *Brit. med. J.*, 1950, ii, 585; Ungley *et al.*, *ibid.*, 905.

The site of formation of intrinsic factor is that part of the gastric mucosa which contains the peptic and oxyntic cells ; these are to be found throughout the mucosa except in a thin strip round the cardiac orifice (cardiac glands)

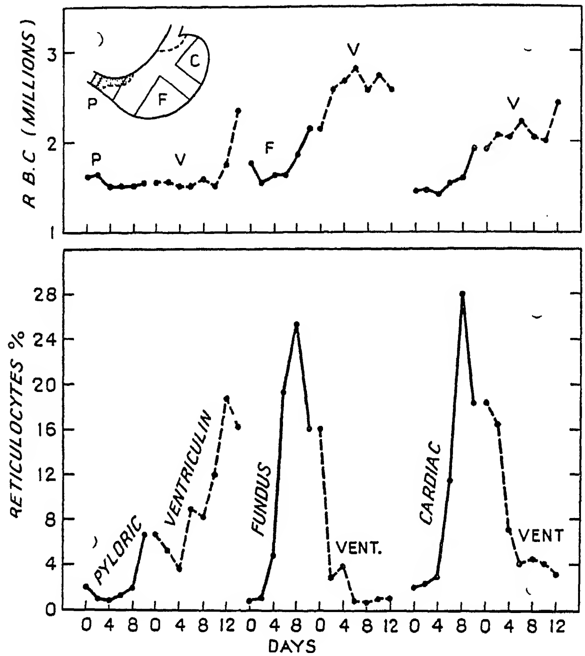


FIG. 113.—Experimental Demonstration of Localization of Intrinsic Factor in " Main Gastric Glands " in Man.

Case of pernicious anæmia. Upper records : red cell count (in millions per c.mm.) ; lower records : reticulocytes per cent. Time in days.
Inset : diagram of human stomach showing distribution of different types of glands (cf. p. 775). Interrupted line near cardiac orifice = distal border of cardiac glands. Interrupted line near pylorus = proximal border of pyloric glands. Rest of stomach = glands containing peptic and oxyntic cells (" main gastric glands ").

Different parts of normal human gastric mucosa were ground up with beef muscle, dried, and administered as shewn (at O). Positive reticulocyte and red cell responses indicate presence of intrinsic factor in region of gastric mucosa examined. Negative responses indicate absence of intrinsic factor. In each case after 8 days' experimental period the response to ventriculin (= hæmatinic principle) was examined.

Note that mucosa from regions C and F (of inset), *i.e.* from the so-called " cardiac " and " fundic " region (both these regions contain " main gastric glands ") is active ; it thus contains intrinsic factor. Mucosa from the pyloric region is only feebly active ; the slight activity is probably due to inclusion of " fundic " mucosa in the material used. (Data from Castle and Fox, *Amer. J. med. Sci.*, 1942, *203*, 18.)

and in part of the pylorus (pyloric glands). Experiments were carried out as follows : the mucosa was stripped off from various regions of the normal human stomach, ground up with ox muscle, dried, and administered to patients with pernicious anæmia. The mucosal area examined which contains

intrinsic factor reacts with the ox muscle to liberate hæmatinic principle. Typical reticulocyte responses were obtained with mucosa from the regions labelled C and F in the stomach map in Fig. 113; a very weak response was obtained with pyloric mucosa but a potent hæmatinic principle gave a characteristic reaction. These experiments demonstrate conclusively that the so-called "main gastric glands" form intrinsic factor. It is exactly this region of the mucosa which is atrophied or destroyed in pernicious anæmia as seen post-mortem, or during life on gastroscopic examination. The common clinical findings of achlorhydria (even after injections of histamine (p. 783)), and sometimes complete achylia, are indicative of a lesion in the same site.

The primary fault in pernicious anæmia is thus the gastric lesion, which results in absence of intrinsic factor; consequently no hæmatinic principle is formed. *Gastrectomy* in man, however, very rarely leads to pernicious anæmia; this result may be explained by the fact that the gastrectomy may have to be total before the body is deprived of all intrinsic factor.

FATE OF HÆMATINIC PRINCIPLE.—The principle, formed in the stomach, passes into the small intestine, is absorbed from there into the portal blood, and so reaches the liver. It is presumably stored there, accounting for its presence in high concentrations in liver extracts. It passes from the liver to the bone marrow where in some unknown way it promotes maturation and normalization of the erythroid cells.

Ventriculin.—An extract of partially autolysed whole stomach wall of the pig—called *ventriculin*—is curative in pernicious anæmia. It has been suggested that the intrinsic factor in the gastric mucosa acts after death on the extrinsic factor in the muscular coat, liberating the hæmatinic principle or a substance of a similar nature.

Causes of Megaloblastic Anæmia.[1] (PERNICIOUS ANÆMIA TYPE).—Theoretically, the hæmatinic principle might not be available for use by the bone marrow in one of five conditions, in all of which a megaloblastic anæmia of the pernicious anæmia type develops characterized by megaloblastic transformation of the marrow erythroid cells and their failure to mature into erythrocytes.

(1) ABSENCE OF INTRINSIC FACTOR.—See *Pernicious Anæmia* (p. 196).

(2) ABSENCE OF EXTRINSIC FACTOR.—In the Tropics, malnutrition is exceedingly common, and megaloblastic anæmia frequently develops. Some cases can be successfully treated by the administration of yeast or yeast products like Marmite, which are rich in extrinsic factor; the essential cause of the disorder is the absence of the extrinsic factor from the diet, which results from its unsatisfactory composition.

(3) FAILURE OF ABSORPTION OF THE HÆMATINIC PRINCIPLE.—This may result from an abnormal state of the intestinal tract in intestinal strictures or faulty anastomoses, impermeability of the intestinal mucosa, or the presence of parasites (like dibothriocephalus latus), which alter or destroy the principle.

[1] *Megaloblastic Anæmia.*—The blood disorders in which there is megaloblastic transformation of the bone marrow and anæmia caused by failure to form or use the hæmatinic principle, are known as megaloblastic anæmias; the circulating red cells are macrocytes, *i.e.* their average diameter is larger than normal. In some cases of malignant disease, cirrhosis of the liver, myxœdema, etc., a macrocytic anæmia may occur associated with the appearance of large normoblasts in the marrow but there are no megaloblasts in the sense defined on p. 196. These anæmias are called *macronormoblastic* anæmias.

Some cases of pernicious anæmia may be thus caused; they do not respond to liver given by mouth but react rapidly to liver extracts given parenterally (*e.g.* intramuscularly).

In the tropical disease, *sprue*, all the above disturbances may be in operation: lack of intrinsic factor in the stomach (with or without achlorhydria), deficiency of the extrinsic factor, and defective absorption of the hæmatinic principle because of the intestinal derangement. Liver extract is effective especially parenterally, but sometimes large doses of iron are needed too.

(4) Disease of the Liver.—In liver damage, pernicious anæmia has been recorded, presumably from failure to store or mobilize the hæmatinic principle.

(5) Cases are described in which anæmia of the pernicious type is uninfluenced by treatment given by any route; it is suggested that the marrow may be unable to utilize the hæmatinic principle supplied to it.

IRON METABOLISM.[1] IRON DEFICIENCY ANÆMIAS

Body Iron.—The body of an adult man contains about 4·5 g. of iron which is distributed between four main forms:

(i) *Blood hæmoglobin*, about 2·5 g. (p. 175).

(ii) *Myohæmoglobin* in varying amounts in *red* muscles.

(iii) *Intracellular enzymes* containing iron-porphyrins, such as cytochrome, cytochrome-oxidase, catalase, and peroxidase (less than 0·1 g.).

(iv) Iron bound with a special tissue protein called *apoferritin* (mol. wt. 465,000) to form the tissue iron-storage compound, *ferritin*. When fully saturated with iron, ferritin may contain 23% of its dry weight as iron. This so-called "storage" iron may amount to 1·5 g. in all.

Myohæmoglobin, peroxidase, and catalase consist of *distinctive proteins* bound with the same porphyrin as is found in the hæm of hæmoglobin, namely *iron-protoporphyrin IX*; their molecular weights are respectively 17,000, 44,000, and 225,000 (cf. hæmoglobin, 68,000). The striking differences in the properties of these compounds must depend on the *protein* part of the molecule. Cytochrome and cytochrome-oxidase contain an iron-protoporphyrin nucleus which differs from iron-protoporphyrin IX in the side chains which are attached to the pyrrol rings; the attached proteins are also different.

The functions of the *tissue iron-porphyrins* (*i.e.* (ii) and (iii) above) can be briefly summarized thus:

(i) *Myohæmoglobin* unites loosely and reversibly with molecular oxygen. Its O_2 dissociation curve resembles that of a CO_2-free hæmoglobin solution, *i.e.* it is a rectangular hyperbola and is shifted well to the left (cf. p. 411); it also resembles fœtal hæmoglobin in its behaviour (p. 1099). It can thus take up O_2 readily even at low O_2 pressures, but only releases it when the tissue O_2 pressure is *very low*. It therefore serves as a small tissue O_2 reserve for the needs of *very* vigorous muscular activity. The properties of red and pale muscle are compared on p. 587; it is curious to note that although red

[1] Granick, *Ann. N.Y. Acad. Sci.*, 1947, *48*, 657. Hahn, *Advances in Biology and Medical Physics*, 1948, *1*, New York.

muscles contain myohæmoglobin, they respond more slowly than pale and consequently develop tetanus at a lower peak tension.

(ii) *Peroxidase* "activates" H_2O_2 to oxidize suitable substrates.

(iii) *Catalase* decomposes H_2O_2 to form water and *molecular oxygen*.

(iv) *Cytochrome* and *cytochrome-oxidase* are concerned with *oxidation processes* in the tissues and *electron transfer*, and not, like hæmoglobin, with the transport of molecular oxygen.

Iron in Food.—It may be deduced from what has been said above that animal foods contain iron in minute amounts which vary with the ferritin and blood content. The approximate iron content in *mg.* per 100 g. of some common foodstuffs is: meat (muscle) 3·5, liver 7·0, egg 2·5, cheese 1·5, herring and haddock 1·0, lentils and peas 5–7, oatmeal 4·0, bread 1·0, golden syrup 1·7, cabbage 1·0, rice and potatoes 0·5. The iron content of milk is low.

Iron Balance.—(1) ADULT MEN.—The intracellular iron-containing enzymes and myohæmoglobin are stable substances the iron of which cannot be called upon for other purposes; blood hæmoglobin on the other hand is continually undergoing destruction. The iron content of hæmoglobin is 0·33%; thus 100 c.c. of blood containing 15 g. of hæmoglobin contain 50 *mg. of iron*. As the red cells live about 100 days, 1% of the total blood hæmoglobin, *i.e.* that contained in 50 *c.c. of blood* is destroyed daily, releasing about 25 *mg. of iron*. It is essential to remember that for all practical purposes iron that has been absorbed from the intestine or some parenteral route *is not subsequently excreted from the body*, or only in negligibly small amounts. Thus an adult man on an adequate iron intake only excretes (in the urine) 0·1 mg. daily. Likewise, hardly any iron is excreted into the bile, and none by the mucosa of the alimentary canal. It follows, therefore, that the iron which is released from the destruction of hæmoglobin, after being temporarily stored in the reticulo-endothelial system (presumably as ferritin) is used again for fresh hæmoglobin synthesis. It will be pointed out later that iron is absorbed from the food in the intestine with considerable difficulty, *i.e.* only a fraction of the food iron is actually absorbed into the body, the rest being *passed out in the fæces*. It is, therefore, hard to say what iron *intake* is needed to provide any particular bodily iron need. Whenever iron loss exceeds iron absorption the blood hæmoglobin falls, *i.e.* anæmia develops. The food iron requirement of a healthy adult male (who is not suffering from any form of blood loss) is negligibly small; a food iron intake of 5 mg. daily is adequate to maintain a normal state of the blood in such healthy men.

(2) ADULT WOMEN.—The iron loss (and therefore the food iron needs) of women is greater because of: (i) the blood loss during the monthly menstrual period; (ii) the iron drain on the mother during pregnancy and labour.

(i) *Menstruation.*—The monthly blood loss in one series of normal women varied between 6 and 180 c.c.; the average was 50 c.c., corresponding to 25 mg. of hæmoglobin iron. A group of women whose monthly loss was 240 c.c. on an average (=120 mg. of iron per month, or 4 mg. daily) developed anæmia, because their food iron intake did not cover their iron loss. Clinical evidence suggests that when the monthly menstrual blood loss exceeds 140 c.c. (=70 mg. of iron per month or 2·3 mg. of iron daily) anæmia develops if the food iron content is below 7 mg. daily (Fig. 114).

by the blood loss during labour. Anæmia commonly develops during pregnancy from iron lack (Figs. 114 and 115).

(3) Growing Children.—During the period of growth the blood must "grow" to keep pace with the rest of the body; a positive iron balance must therefore be maintained to enable the necessary additional hæmoglobin to be formed. The following data indicate the size of the problem. At birth a baby weighs 3 kg. (7 lb.) and contains 300 c.c. of blood; the total iron content of the body is 400 mg. At one year it weighs 10 kg. (21 lb.) and contains about 1000 c.c. of blood. To produce the necessary increase in hæmoglobin and body iron a positive iron balance of 500 mg. of iron is required (=1·3 mg. daily) over and above the iron needed to compensate for normal wastage. Mackay's data (Fig. 100) show that even when mother and infant are on a satisfactory diet the infant's hæmoglobin level falls from a birth value exceeding 100%, to 75% at 3 months; between 6 and 12 months it is 85%. Hæmoglobin formation cannot, therefore, even under good circumstances keep pace with the increase in plasma volume. In families on a low iron intake the infant is probably born with poor iron reserves on which to draw, and it gets less iron in its diet; as a result the hæmoglobin falls progressively, reaching a minimum at the end of the first year (Fig. 114). Subsequently, as the rate of growth slows down, the blood picture improves steadily during childhood. A fresh strain develops in girls with the onset of puberty and menstruation.

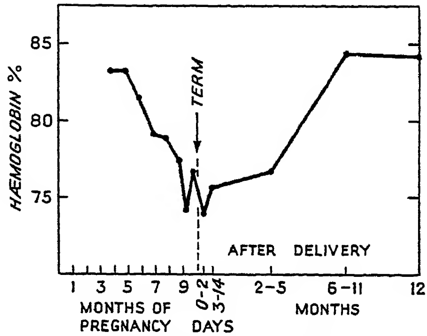


Fig. 115.—Development of Anæmia in Women during Pregnancy and Progressive Recovery after Delivery. (Modified from Davidson and Fullerton, *Edin. med. J.*, 1938, *45*, 9.)

Ordinate: hæmoglobin %.

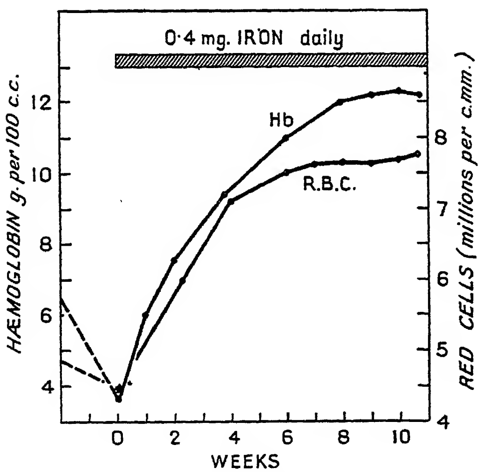


Fig. 116.—Results of Iron Deficiency in Young Rats. (Tyson *et al.*, *Amer. J. clin. Path.* 1939, *9*, 63.)

The offspring of iron-deficient mothers were weaned at 3 weeks and put on a milk diet. Untreated rats died when the hæmoglobin fell to 3 g-%. On adding iron (at 0) to diet (0·4 mg. daily) rapid recovery occurred.

Ordinate: hæmoglobin in g. per 100 cc.; red cells in millions per c.mm.

The problem can be studied experimentally in rats. The mothers were

kept on an iron-free diet and the offspring after weaning were fed on milk. The blood hæmoglobin concentration steadily fell because while the animal was growing and increasing its plasma volume it was not making fresh hæmoglobin (Fig. 116). If the young rats or the anæmic children are given adequate amounts of iron, rapid recovery of the blood occurs.

Blood Iron.—Iron is present in the blood in two forms: (i) as *hæmoglobin* in the red blood cells; 15 g. of hæmoglobin per 100 c.c. blood corresponds to 50 mg. of inorganic Fe; (ii) as *serum iron* amounting to 0·05–0·18 mg–% (=50–180 μg–%); this is the form in which iron transport takes place. Serum iron is in the *ferrous* form and combined with a serum protein (β-globulin). Substantial changes in serum iron take place after ingestion of therapeutic doses of iron salts (p. 209); but under ordinary conditions of food intake the fluctuations in serum iron are slight.

Iron Absorption.—(1) Food Iron.—Very little is known about the changes which food iron undergoes in the alimentary canal. It is believed that inorganic iron is split off from ferritin, but that the digestive juices cannot release the iron which is bound with the porphyrin molecule in the hæm compounds of food. Food may contain traces of inorganic iron salts if rusty knives or chipped enamel saucepans are used in the preparation and cooking of food; these may be a more important source of iron than is generally supposed. Iron is absorbed in the *ferrous* form possibly in the stomach and certainly in the duodenum and upper small intestine; the presence of bile salts does not promote iron absorption. The iron is taken up by the cells of the intestinal mucosa and combined with apoferritin there, to form ferritin, which releases its contained iron into the circulation as and when required for hæmoglobin formation.

In normal adult males iron absorption is very small, as shown by the following iron balance experiments.

I. Low Iron Intake (all data in mg. for 14-day periods):

Total iron in food . . .	100·5 mg. (7·1 mg. per day).
„ „ urine . . .	1·3 mg. (0·1 mg. „).
„ „ fæces . . .	102·5 mg. (7.3 mg. „).

In this experiment no iron was absorbed; all the ingested food iron was passed out in the fæces; the excess loss in the fæces may have been due to shedding of lining cells of the intestinal mucosa which contain iron.

II. High Iron Intake (all data in mg. for 14 days):

Iron intake as food . . .	110·6 mg.
„ „ medicinal iron .	82·8 mg.
Total Intake . .	193·4 mg. (13·8 mg. per day)

Iron loss in urine=1·5 mg. (0·1 mg. per day).
Iron loss in fæces=187·3 mg. (13·4 mg. per day).
Positive iron balance=4·6 mg. (0·3 mg. per day).

In this experiment only 0·3 or 0·3+0·1 mg. of iron were absorbed daily.

It is found clinically that a total iron food intake of 10–15 mg. daily is sufficient to maintain a normal state of the blood in all physiological states in women and children and is well above the minimal requirements of normal men.

(2) Therapeutically administered Iron Salts.—In cases of iron deficiency anæmia, iron is administered in the form of inorganic salts. The fate of the iron can be readily followed if it is administered in the radio-active form as $^{59}Fe^*$; its half-life is 47 days. Fig. 117 shows the effects of administering 84 mg. of radio-Fe in a single dose by mouth to a dog. Absorption sets in rapidly as proved by the appearance of radio-Fe in the serum within one hour; serum radio-Fe reaches its peak at 12 hours and falls to normal in 24 hours. Within 2 hours of ingesting the radio-Fe it begins to appear in the circulating corpuscles, proving that it has reached the bone marrow and has been used there in the manufacture of fresh hæmoglobin; the radio-Fe has been added to the body's general iron reserve, *i.e.* it has entered the body's

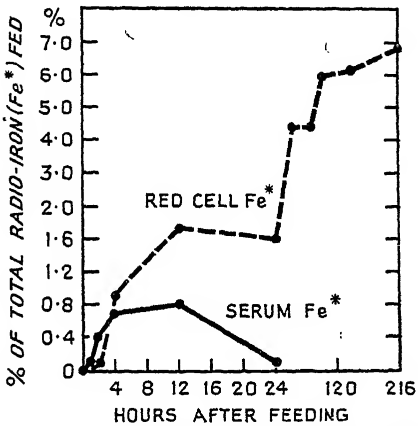


Fig. 117.—Absorption of Radio-Iron (Fe*) from the Intestine and its Incorporation in Red Cells.

Ordinate: percentage of total radio-iron (Fe*) given by mouth found in plasma or red cells. Dogs given single dose of 84 mg. of radio-iron by mouth. (Whipple *et al.*, *J. exp. Med.*, 1939, *69*, 747.)

" iron pool." No distinction is subsequently drawn by the body between the newly absorbed iron and that previously present in the tissues.

The following factors influence the degree of absorption of iron salts:

(i) Ferrous (Fe^{++}) salts are more readily absorbed than ferric (Fe^{+++}) salts; in different experiments the ratio obtained was 1·5–10·0 of Fe^{++} to 1·0 of Fe^{+++}. (Fig. 118.)

(ii) Absorption is independent of the secretion of HCl by the stomach.

(iii) Insoluble iron salts (*e.g.* ferrous phosphate) are not absorbed.

(iv) The absolute amount of any particular salt which is absorbed depends on the dose. With minute doses the *percentage* absorption is high but the *absolute* amount absorbed is still minute; as the dose is increased the percentage absorption progressively falls but the absolute amount absorbed is increased. Thus the total iron content or the iron concentration in the bowel is an important factor in determining the total amount absorbed; and it is the *absolute* amount of iron absorbed which really matters.

(v) With a given iron intake, a larger amount is absorbed in *anæmic* subjects than in normals. This is well shown in Fig. 119.

The explanation is obscure. The rate of iron absorption may depend on the degree of saturation with iron of the apoferritin of the intestinal mucosa. In normal subjects the degree of saturation is high and the iron uptake is, consequently, small. In anæmic states as the tissue "reserve" iron is drawn upon, food iron is more readily taken up by the relatively *unsaturated* apoferritin in the intestinal mucosa. It is claimed also that when the iron intake is large, more apoferritin is synthesized in the intestinal mucosa, thus facilitating further iron uptake.

Intravenously injected Iron Salts.—Iron thus injected is not excreted by the intestinal mucosa from the blood into the lumen of the gut; it is, therefore, retained in the body, in the macrophages of the liver and spleen, throughout the spleen pulp, in the hepatic cells and, in traces, in the stomach and large intestine. There is still uncertainty whether such stored iron can be called upon for hæmoglobin formation; it can only be eliminated from the body very slowly, if at all.

Factors controlling Hæmoglobin Formation.[1]—(i) *Quantitative Technique.*—Whipple's *standard anæmic dog* is employed. A normal dog weighing 10 kg. has about 200 g. of hæmoglobin in the circulation. The animal is bled weekly to reduce the total circulating hæmoglobin each time to 60 g. During the first 6–8 weeks the dog seems to be able to mobilize unidentified reserves of hæmoglobin to help to make good the deficiency, but these are finally exhausted. Subsequently the amount of hæmoglobin newly formed each 14 days is readily measured by determining the amount of hæmoglobin that has to be removed to bring the total in the blood down to the basal level of 60 g. These studies show that in addition to iron, dietary *protein* is of great importance in the treatment of hæmorrhagic anæmia.

(ii) *Rôle of Protein.*—A low protein intake retards hæmoglobin regeneration even in the presence of excess iron; the limiting factor here is lack of *globin.* Globin itself (in the food) is the protein which is used most economically in hæmoglobin formation. Some food proteins are less effective than others; thus bread and other cereals, dairy products, most vegetables and fruits, and salmon are relatively inert. Liver, kidney, spleen, and heart are most potent, muscle occupying an intermediate position. Though liver is of great value in the treatment of hæmorrhagic anæmia in dogs, it seems to have *no* such exceptional efficacy in hæmorrhagic anæmia in man. Studies with amino-acid mixtures show that the usual essential amino-acids are needed for hæmoglobin formation.

(iii) *Rôle of Porphyrins.*—The porphyrin content of the diet is unimportant as it is readily synthesized.

(iv) *Rôle of Iron.*—See pp. 205 *et seq.*

(v) *Rôle of Bile.*—If a *bile fistula* is established so that all the bile escapes to the exterior, the rate of hæmoglobin formation is halved; the administration of bile daily by mouth is, however, without effect. Iron salts given by mouth to these animals produce half the expected new hæmoglobin formation. The mode of action of a bile fistula is obscure: it is associated with impaired liver function and diminished globin formation.

[1] Whipple, *Amer. J. med. Sci.*, 1938, *196*, 609; Whipple *et al.*, *J. exp. Med.*, 1945, *82*, 311; *ibid.*, 1947, *85*, 243.

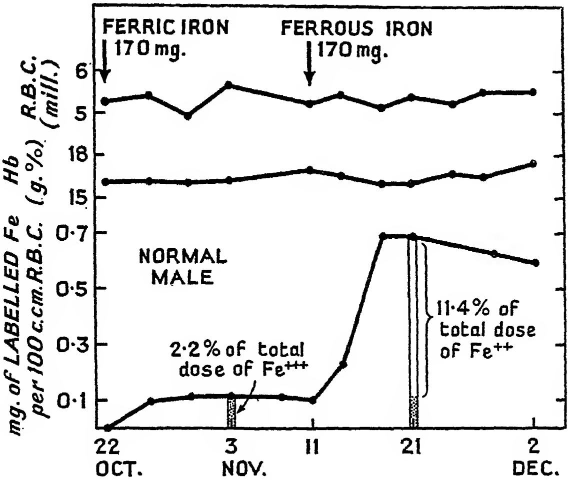


Fig. 118

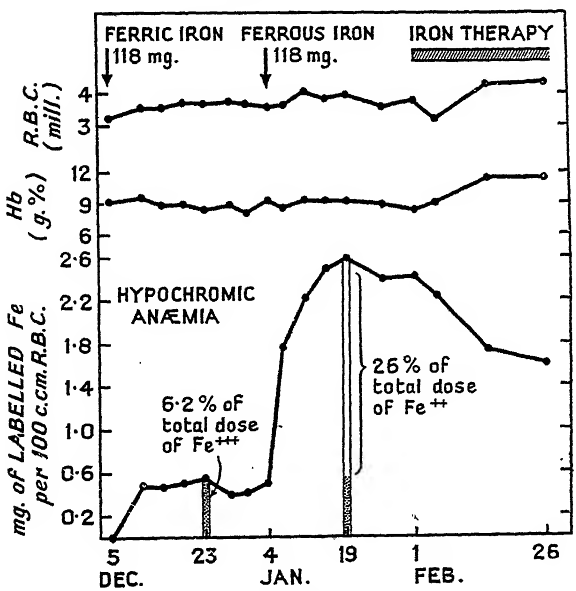


Fig. 119

Figs. 118 and 119.—Absorption of Radio-active Iron in Normal Person and in Patient with Hypochromic Microcytic Anæmia. Ferrous iron is more readily absorbed than Ferric Iron. Iron is more readily absorbed in Anæmic Patients than in Normal Persons. (Moore *et al.*, *J. clin. Investig.*, 1944, *23*, 762, 764.)

	Normal. Fig. 118.	*Hypochromic Anæmia.* Fig. 119.
Initial red cell count per c.mm.	5·2	3·2
Initial hæmoglobin g./100 cc.	16·0	9·0
Amount of radio-Fe^{+++} given by mouth as $FeCl_3$	170 mg.	118 mg.
Fraction of total dose of radio-Fe^{+++} found in circulating red cells	2·2%	6·2%
Amount of radio-Fe^{++} given by mouth as $FeCl_2$	170 mg.	118 mg.
Fraction of total dose of radio-Fe^{++} found in circulating red cells	11·4%	26·0%

Labelled Fe=radio-active iron.

At the beginning of February the anæmic patient was given intensive iron therapy; the red cell count and hæmoglobin concentration rose in about 4 weeks to 4·5 million and 12 g–% respectively.

Iron Deficiency Anæmia.—Iron deficiency anæmias are of the *microcytic hypochromic* type (p. 168); the cell diameter and volume are below average and the cells are poorly filled with hæmoglobin. Both red cell count and hæmoglobin values are lowered, the latter to a greater degree than the former. The marrow may show *normoblastic* hyperplasia.

As is evident from the discussion on pp. 205 *et seq.*, iron deficiency anæmia occurs: (i) in infants and young children; (ii) in women during active reproductive life; (iii) in anyone who suffers severe blood loss. The co-existence of factors which inhibit hæmoglobin or red cell formation aggravates the anæmia. The following complications commonly occur:

(i) The presence of *infection*:[1] this is especially important in infants, in whom an attack of bronchitis, otitis media, or pyrexia from any cause may lower the blood hæmoglobin level, sometimes even when iron is being medicinally administered at the same time (Fig. 120).

(ii) An unsatisfactory diet, especially one which is lacking in (*a*) *protein*, which is needed for globin formation, or (*b*) *vitamins*: the lack of vitamins

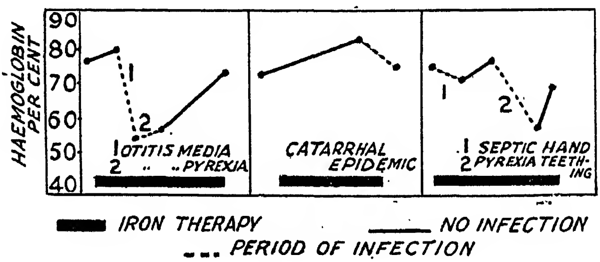


FIG. 120.—Effect of Infection on Hæmoglobin Level in Infants. (Davidson and Fullerton *Edin. med. J.*, 1938, *45*, 12.)

may lead to an unhealthy state of the alimentary canal which depresses absorption of the foodstuffs and may also impair the tissue mechanisms which are involved in hæmoglobin synthesis. *Achlorhydria* is said to interfere with the proper absorption of *food* iron (though it does not affect the absorption of therapeutic iron salts).

(iii) In pregnancy the formation of intrinsic factor may be impaired, leading to development of a macrocytic anæmia (Fig. 115). In addition, the blood hæmoglobin concentration may be lowered because of an increase in plasma volume.

Pure iron deficiency anæmia is cured by giving adequate doses of suitable iron salts. These doses are enormous compared with the normal iron intake in the food (Fig. 121, A). The amount of medicinal iron *utilized* for fresh hæmoglobin formation can be readily calculated from the rate of rise of the *total* blood hæmoglobin content which optimally is 1·5 g. (=5 mg. iron) daily. Such a rate of blood improvement may occur on giving 1200 mg. (90 grains) of $Fe(NH_4)$Citrate daily; the iron utilized in this instance is under 1% of the iron administered. The *optimum* utilization rate with $Fe(NH_4)$ Citrate is 3% compared with 20% for ferrous sulphate.

Iron therapy acts not only by supplying an essential raw material but it

[1] Wintrobe *et al.*, *J. clin. Investig.*, 1947, *26*, 103, 114, 121.

also *stimulates* the red marrow to make fresh cells and hæmoglobin. Even in normal subjects iron medication may increase the blood hæmoglobin level, but this increase is transient, the blood returning to its original state in spite of continued treatment (Fig. 121, B).

ADJUVANT ACTION OF COPPER.[1]—In mammals copper is related to blood formation.[2] Iron-deficiency anæmia in rats (Fig. 116) is not cured by *absolutely pure* iron salts. If 0·05 mg. of copper, together with 0·5 mg. of iron are added to the daily ration of milk, rapid cure occurs. A normal mixed diet contains adequate amounts of copper. Most medicinal iron salts contain some admixture of copper, but less than the optimum for the cure of anæmia, *e.g.* only 0·02 mg. of copper per g. of iron salt.

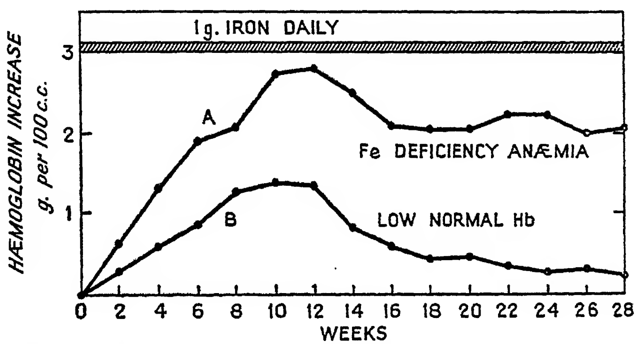


FIG. 121.—Effects of Iron Administration on Blood Hæmoglobin Concentration in Man

In both A and B, 1 g. of iron was administered daily throughout the 28-week period. Ordinate represents *increase* in hæmoglobin concentration *per 100 c.c. of blood*. The *total* increase in blood hæmoglobin in g. is obtained by multiplying the value on the ordinate by $\frac{\text{blood volume}}{100}$, *e.g.* $\frac{5000}{100}=50$.

A. Results in group of cases of *iron deficiency anæmia*. Maximal improvement occurs after 10 weeks there is a *small* subsequent decline.

B. Results in group of subjects with *low normal* hæmoglobin. There is a smaller initial rise n hæmoglobin concentration; then in spite of continued iron administration the hæmoglobin concentration falls to its *initial value*. (Fowler and Barer, *Amer. J. med. Sci.*, 1941, *201*, 648.)

The mode of action of copper is unknown; it may be related to the building up of iron into hæmoglobin.

Some clinicians claim that better results are obtained in cases of hypochromic anæmia if copper is given together with iron (in the ratio of copper to iron of 1 to 100) than with iron therapy alone.

Hæmochromatosis.[3]—In this disease there is an excessive deposition of iron in all the organs (especially in the liver and pancreas), as brown granules

[1] Elvehjem, *Physiol. Rev.*, 1935, *15*, 471.

[2] Copper is widely distributed in organic matter and is found in minute traces in sea water, as is shown by the following data: Sea water, 1×10^{-11}; plants, $3\text{–}40\times10^{-6}$; blood of marine animals, $3\text{–}23\times10^{-5}$; brain, $3\text{–}6\times10^{-6}$; milk, $2\text{–}5\times10^{-7}$. The blood of marine animals contains an oxygen-carrying pigment called *hæmocyanin*, which contains 0·38% of copper; the copper acts in the same way as the iron in hæm, and there is a quantitative relationship between the amount of copper and the oxygen-carrying power. A highly coloured pigment (*turacin*) has been obtained from the feathers of a South African bird, the turaco, and contains 7% of copper; turacin is a copper-porphyrin (cf. p. 174).

[3] Sheldon, *Hæmochromatosis*, London, 1935. Granick, *Bull. N.Y. Acad. Med.*, 1949, *52*, 403.

of *hæmosiderin*, a compound resembling *ferritin* (p. 204); as a result the tissue cells are damaged and undergo fibrosis. An iron-free pigment—*hæmofuscin*—is also deposited. The pathological accumulation of iron may sometimes be due to a primary abnormality of the intestinal mucosa, leading to excessive iron absorption from the gut. In one patient with hæmolytic anæmia who received about 200 blood transfusions, the released iron was deposited in the tissues, causing fatal hæmochromatosis.

THE WHITE BLOOD CORPUSCLES

The white blood corpuscles (leucocytes) are divided into three groups:

(1) *Granulocytes* (10–14 μ) characterized by the presence of granules in the cytoplasm, and a lobed nucleus. Using Leishman's stain, three types of cell can be recognized by the character of their granules: *neutrophil* (or polymorphonuclear leucocytes) with fine red-brown granules; *eosinophil*, crammed with large red granules; and *basophil* containing purple-blue granules. The nuclear chromatin of this group of leucocytes is coarse and "ropy." (Plate III.)

(2) *Lymphocytes* (small, 7–10 μ; large, 10–14 μ): these are round non-granular cells with large round nuclei which practically fill the cell substance. They are divided into *large* and *small* lymphocytes. The nuclear chromatin is coarser and "lumpy." (Plate IV.)

(3) *Monocytes* (10–18 μ): this is a convenient term used to describe a group of cells which are ill understood and suffer from a multitude of labels. The group includes the mononuclear cells, hyaline, and transitional cells among others. They are large pale cells with a pale-staining round or indented eccentric nucleus, the chromatin of which is finely reticular and non-granular protoplasm. Representative monocytes are depicted in Plate IV.

Normal Count.[1]—The normal range of the white cell count is 4000–11,000 per c.mm. Considerable variations between these limits may occur in the same individual from day to day, from hour to hour, and even from minute to minute. The count is made up approximately as follows: granulocytes 70% (neutrophils 50–70, eosinophils 1–4, basophils 0–1), lymphocytes 20–40, monocytes 2–8%. (In children neutrophils are about 10% less and lymphocytes about 10% more than in adults.) Representative *absolute* numbers of the different cells per c.mm. might be as follows: neutrophils 3000–6000, eosinophils 150–400, basophils 0–100, lymphocytes 1500–2700, monocytes 350–800.

An increase in the white count may be produced by exercise or the injection of adrenaline. The leucocytosis alleged to be induced by meals probably represents nothing more than the normal random fluctuation. The white count rises during the latter months of pregnancy and reaches its peak (about 17,000 per c.mm.) with the onset of labour.[2]

It may be supposed that stores of white cells exist in various parts of the body, and they may appear and disappear from the circulation with great rapidity under various circumstances.

[1] Garrey and Bryan, *Physiol. Rev.*, 1935, *15*, 597.

[2] Nordenson, *Quart. J. Med.*, 1939, *8*, 311. Sturgis and Bethell, *Physiol. Rev.*, 1943, *23*, 279.

In new-born infants the white count is about 20,000 per c.mm.; after the second week it begins to decline (Fig. 122). During infancy the lymphocyte is the predominant blood cell, constituting 40–50% of the count.

Development of White Corpuscles—In the embryo the white corpuscles develop in the mesoderm and migrate secondarily into the blood vessels. In extra-uterine life the granulocytes develop normally exclusively in the *red marrow*; the lymphocytes and monocytes also develop to a slight extent in the marrow, but their main site of origin is in the *lymphoid tissues* of the body (Plates III and IV).

Granulopoiesis.[1]—The process can be well studied in the artificially simplified marrow of a pigeon or rabbit (p. 163). Myeloid activity is best seen

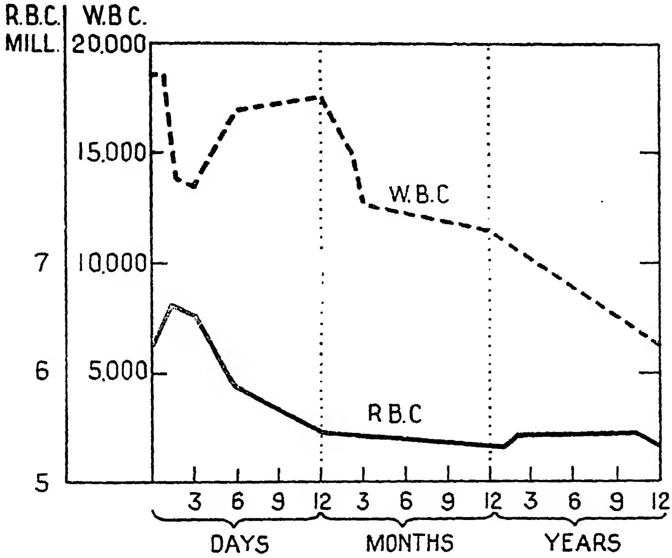


FIG. 122.—Red and White Cell Count from Birth to 12 years of age. (Whitby and Britton, *Disorders of the Blood*, Churchill.)

at the periphery of the marrow in close proximity to *dilated* sinuses, and is wholly an *extra-vascular* process (Fig. 123). Between the fat cells, branched *reticulum cells* can be seen—irregular in outline, free from granules or mitochondria, and with faintly basophil cytoplasm. These cells multiply by mitosis to give rise to the *primitive white blood cell*: the nucleus is large and spherical and surrounded by a thin rim of basophil protoplasm. The stages in the development of each type of granulocyte are: primitive white cell, myeloblast, myelocyte A or premyelocyte, myelocyte B or myelocyte proper, myelocyte C or metamyelocyte, leucocyte. These cells may all be identified in stained marrow smears obtained from sternal punctures in man and they show the typical, successive changes of maturity (p. 164).

[1] Doan, Cunningham, and Sabin, *Contribution to Embryology*, Carnegie Publications, 1925, *16*, 227. Blackfan, Diamond, and Leister, *Atlas of Blood in Children*, N.Y., 1944.

Myeloblast (12–18 μ).—This cell develops from the primitive white cell; the nucleus is pale purple-blue, large and round with finely stippled chromatin and several nucleoli. The cytoplasm consists of a narrow blue rim without granules.

FIG. 123.—Granulopoiesis in Bone Marrow.

Marrow of Pigeon. Zone of granulopoiesis 24 hours after feeding following hypoplasia induced by starvation. Vessels injected with Indian ink. Practically every vessel in the field is widely dilated and well filled with ink. In the intervascular spaces are numerous *myelocytes*. (Doan, Cunningham, and Sabin.)

Myelocyte (12–18 μ).—These cells are characterized by the appearance of *granules* in the cytoplasm. Using Leishman's stain, the myelocytes may be classified according to the colour of their granules into *neutrophil* (most), *eosinophil*, and *basophil* (very scanty). The cytoplasm of the myelocytes as a whole is more extensive and less basophilic; the nucleus is smaller and more basophilic, the nucleoli have disappeared and the chromatin is coarser.[1] (See Plate III.)

If vital staining is employed (Janus green and neutral red) three stages in the development of each *variety of myelocyte* may be recognized—types A, B, C, according to the number of granules which are present. Type A (the youngest, or

[1] Chromatin detail (usually ignored in histology classes) gives hæmatologists valuable help in identifying atypical cells.

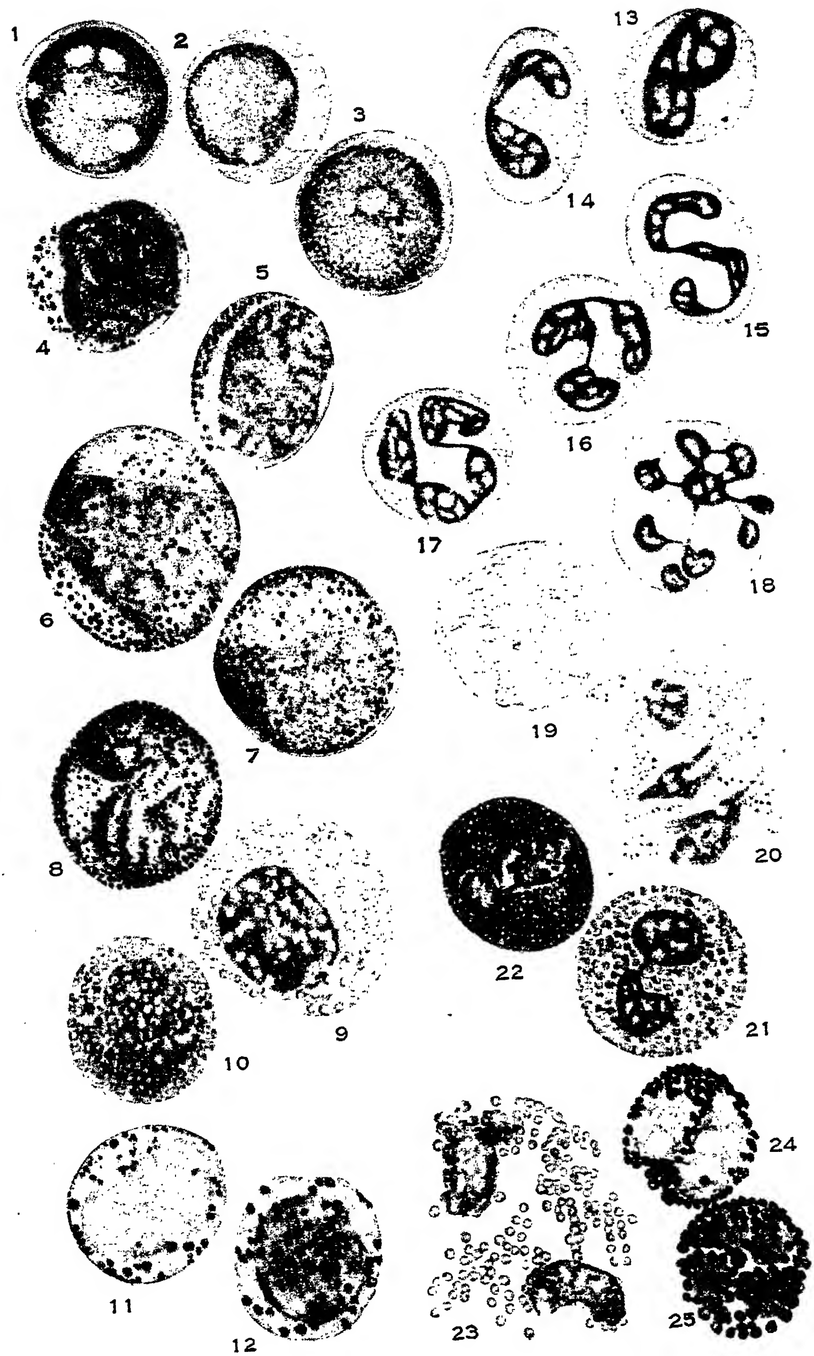
KEY TO PLATE III.

MATURATION IN MYELOID SERIES.

1–3. *Myeloblasts.*
4. *Myelocytes type "A" or premyelocytes.*
5–7. *Myelocytes type "B."*
8. *Myelocytes type "C," metamyelocytes* or *late myelocytes.*
9, 10. *Eosinophil myelocytes.*
11, 12. *Basophil myelocytes.*
13. *Young neutrophil:* also called *early, band-form,* or *non-segmented neutrophil:* nucleus is condensed and beginning to lobulate.
14. *Adult neutrophil:* nucleus has two lobes.
15, 16. *Adult neutrophils:* nucleus has three lobes.
17. *Adult neutrophil:* nucleus has four or more lobes.
18. *Aging neutrophil:* nucleus has multiple lobulation. Cytoplasm contains sparser granules and some vacuoles.
19, 20. *Degenerated neutrophils:* old and fragile cells disrupted in smearing process.
21, 22. *Adult eosinophils:* nucleus usually has two or three lobes; is filled with typical granules.
23. *Aging eosinophil* · old and fragile cell disrupted in smearing process.
24, 25. *Basophils:* nucleus usually has two or three lobes. Cytoplasm is generally filled with very large blue-black granules.

Reproduced from Plates II and III, Blackfan, Diamond, and Leister, *Atlas of the Blood,* The Commonwealth Fund, New York, 1944.

PLATE III



premyelocyte) contains a few granules staining with neutral red and many mitochondria; type C (the oldest, or *metamyelocyte*) contains many granules staining with neutral red and few mitochondria; type B (the *myelocyte proper*) is intermediate. Cells may similarly be classified in smears stained with Leishman.

Leucocytes.—Each type of metamyelocyte gives rise to the corresponding *leucocyte* (neutrophil, eosinophil, and basophil); the nucleus indents, and then becomes lobed; the granules, in fresh preparations (on a warm stage) instead of being motionless, show dancing or streaming movements, and the cytoplasm becomes more liquid, so that *amœboid movements* occur. The more mature white cells are found lying just external to the sinusoids, and *pass actively through the intact endothelial lining* of these vessels into the circulation. *Senile* leucocytes lose their motility, the granules are still and no longer stain with neutral red; the cells break up readily in films, and the nucleus can be seen lying free, surrounded by faintly staining granules. As a leucocyte ages the complexity of the nuclear lobulation increases; thus a very young leucocyte has a horseshoe-shaped nucleus, while an old cell may show four or five lobes joined together by very faint strands of chromatin.

According to Sabin, " showers " of senile granulocytes can be seen several times a day in the circulating blood, when they may constitute 30% of the granulocytes. As the senile cells probably undergo rapid destruction, these findings suggest that the granulocytes function for only a day or two in the blood stream. This may be correlated with the fact that 75% *of the cells of normal marrow are myeloid* and only 25% are erythroid, although the ratio of white to red cells in the *blood* is about 1 : 1000. In regenerating *animal* marrow the percentage of cells per 100 myeloid cells counted, is as follows: neutrophil myelocytes—type C, 86; type B, 4; type A, 0·4; eosinophil myelocytes, 2·5; basophil myelocytes, 1; myeloblasts, 1; polymorphs, 5. It will be noticed that the *predominant cell is the most mature, type C myelocyte* (*metamyelocyte*), *i.e.* the immediate precursor of the mature leucocyte. With sternal marrow smears in *man*, the proportion of myelocytes is lower and that of mature neutrophils higher, owing to admixture with peripheral blood.

In *aplastic anæmia* atrophy of the myeloid (as well as the erythroid) tissue in the red marrow usually occurs and the granulocyte count in the circulating blood is very low. If a large transfusion is given the granulocyte count rises temporarily, but returns to its previous low level in less than 12 hours. This observation supports the conclusions of Sabin about the short life of the granulocytes. If blood is kept in a vessel outside the body the granulocytes die in a few hours, as is proved on examination of the white cells on the warm stage. If these dead cells are transfused they are removed almost instantly from the circulation.

Under certain circumstances, the reticulum cells in the *spleen*, instead of giving rise to lymphocytes or monocytes, may give rise to the myelocyte series of cells (p. 225).

The Lymphocytes.—Development.—Normally the lymphocytes are formed chiefly in the lymphoid tissues of the body, *i.e.* lymph nodes, spleen, thymus, tonsils, Peyer's patches [aggregated lymphatic nodules] in the intestine, and only to a minor extent in the bone marrow. In conditions associated with lymphocytosis and in lymphatic leukemia the marrow

lymphocytes become numerous and overshadow the other types of white cells.

The details of the development of lymphocytes in lymphoid tissue are as follows. In the centre of the lymphatic nodules is the *germinal centre*, consisting of reticulum cells some of which differentiate and proliferate to form large cells with pale nuclei called *lymphoblasts* (Plate IV). These are indistinguishable histologically from myeloblasts. They divide to form *large lymphocytes*, and they in turn become condensed in size to form *small lymphocytes*. The lymphocytes form a more deeply staining zone at the periphery of the lymphatic nodules. They leave the gland in the efferent lymphatics and ultimately enter the circulation via the thoracic and right lymphatics ducts. If the thoracic duct is brought out to open on the surface, there is a substantial diminution in the number of circulating blood lymphocytes. The total number of lymphocytes entering the thoracic duct in the course of a day is about two and a half times as great as the total number present in the blood stream at any one time.[1] This suggests that the life of the lymphocytes in the circulation does not exceed a few hours. The lymphocytes formed in the bone marrow pass directly into the circulation.

FATE AND FUNCTIONS OF LYMPHOCYTES.—There is much doubt about the site of destruction of the lymphocytes; one view is as follows. The germinal centres of the lymphoid nodules are not only concerned with lymphocyte formation but also contain many macrophages filled with fragments of dead lymphocytes; it is suggested, therefore, that the lymphocytes

[1] Yoffey, *J. Anat.*, 1933, *67*, 250.

KEY TO PLATE IV.

MATURATION IN LYMPHOID, MONOCYTE, AND THROMBOCYTE SERIES.

1. *Lymphoblast.*
2. *Large* or *young lymphocyte*: cytoplasm is fairly clear blue.
3. *Large (adult) lymphocyte*: cytoplasm is fairly abundant and is clear light blue, containing several to many large reddish granules towards the periphery of the cell.
4. *Medium lymphocyte.*
5. *Small* or *adult lymphocyte.*
6. *Erythrocyte,* for comparison of size.

7, 8. *Large lymphocytes* of the type most frequently in infectious mononucleosis. Nucleus is dense with some fenestration of chromatin, and is often eccentrically placed. Cytoplasm is abundant, with irregular edge; it is clear light blue, denser at the margin.

9, 10. *Lymphocytes* of plasma-cell type. Nucleus is eccentrically placed, darkly staining with dense chromatic clumps often in cartwheel arrangement. Cytoplasm is abundant, dark blue with clear areas near nucleus, and stains more densely at the periphery; it often has a foamy appearance.

11. *Monoblast.*

12–15. *Monocytes.* Occasionally there are vacuoles, evidence of active phagocytosis, at the periphery of the cytoplasm (as in 15).

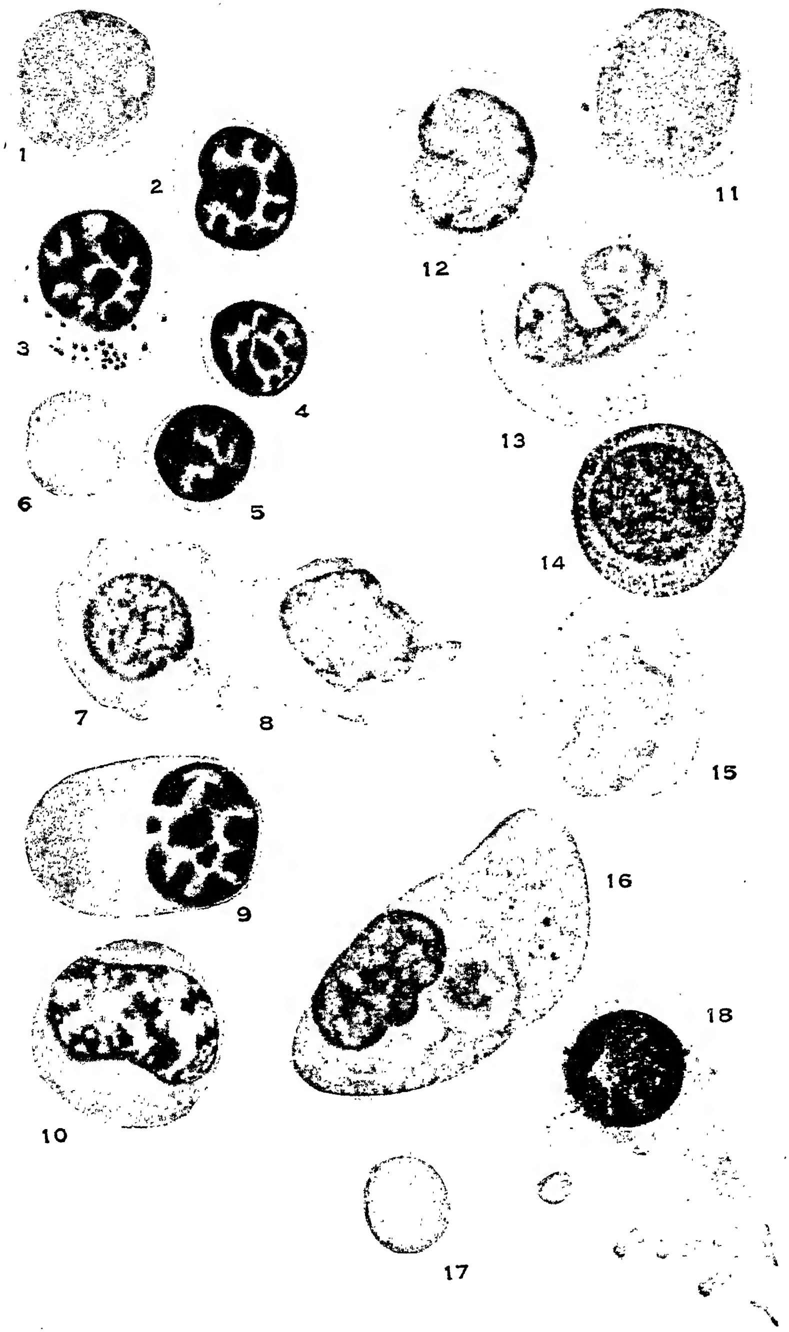
16. *Mononuclear phagocytes, endothelial phagocytes* or *clasmatocytes,* the largest cells ordinarily found in the peripheral blood. Cytoplasm is finely granulated and contains many vacuoles of all sizes, often filled with ingested material.

17. *Erythrocyte,* for comparison of size.

18. *Megakaryocyte* and *Platelets (thrombocytes).* Cytoplasm is light blue with dark purple masses of granules and no visible cellular membrane; it is here seen to trail in long pseudopodia from which masses of platelets break off and are scattered as individual platelets.

Reproduced from Plates IV and V, Blackfan, Diamond, and Leister, *Atlas of the Blood,* The Commonwealth Fund, New York, 1944.

PLATE IV



[Facing page 213

return via the blood (or the lymphatics) to the lymphoid tissues where they are destroyed.

Extracts of lymphocytes contain γ-globulin, which is the plasma protein fraction containing the immune bodies, *i.e.* the substances concerned with protection against infections (p. 137). The dissolution of the lymphocytes thus discharges immune bodies into the efferent lymphatics and so into the circulation. The lymphoid tissues hypertrophy in the conditions enumerated on p. 1017; they atrophy early and markedly in starvation.

Monocytes.—A separate line of development of the primitive white cell in the marrow and in lymphoid tissues gives rise successively to *monoblasts* and *monocytes* (Plate IV). Premyelocytes may closely resemble monocytes but can be distinguished by the peroxidase reaction. The myeloid family in general contains a *peroxidase* which reacts with hydrogen peroxide and benzidine, when these are poured on to a film, to give blue cytoplasmic granules; the enzyme is absent from the monocytes which therefore show no such granules. It must be noted, however, that even in the case of granulocyte series the enzyme does not appear *before* the premyelocyte stage; the peroxidase stain cannot therefore distinguish between myeloblasts and monocytes or between myeloblasts and lymphoblasts.

The circulating monocytes are actively phagocytic and can ingest foreign particles such as pigment and blood protozoa; they are possibly destroyed by the macrophages of the tissues.

Changes in White Cell Count in Disease.—An increase in the total circulating leucocytes above 11,000 per c.mm. is known as *leucocytosis*; a decrease below 4000 per c.mm. is a *leucopenia*. A differential leucocyte count enables the percentage and absolute numbers of the different varieties of white cell to be determined. The *absolute* figures are far more important than the alterations in the relative proportions. According to the type of cell involved, a leucocytosis may be described as neutrophil, eosinophil, or basophil leucocytosis, lymphocytosis, or monocytosis. Leucopenia is generally due to a neutropenia, *i.e.* a decrease in the neutrophil cells.

NEUTROPHIL LEUCOCYTOSIS occurs in many inflammatory conditions, for example, in pneumonia, endocarditis, and pyogenic infections. The neutrophils may number 20,000 or 40,000 or even 150,000 per c.mm., and constitute 90–95% of the total white count. Great multiplication of the myelocytes occurs in the marrow, and leucocytes are discharged into the blood stream in enormous numbers. There is naturally an increase in the proportion of *young* leucocytes in the circulation; this is known as "a shift to the left."

EOSINOPHILIC LEUCOCYTOSIS occurs in infections with parasitic worms, in allergic conditions (hay fever, asthma), scarlet fever, and many skin diseases. The marrow shows great proliferation of the eosinophil myelocytes. The level of circulating eosinophils is regulated directly by the *adrenal corticoids* and indirectly by ACTH. This subject is fully discussed on pp. 951 *et seq.*

ABSOLUTE LYMPHOCYTOSIS is comparatively rare; the *relative* number or *percentage* of lymphocytes is of course decreased in any condition of true granulocytosis and it is increased when there is an absolute decrease in the granulocytes, as in enteric infections and many chronic anæmias. Beyond indicating a neutropenia these percentage changes are of no significance because the total lymphocyte count is unaltered. *Absolute* lymphocytosis is

a normal phenomenon in infants and occurs in certain diseases like whooping-cough and glandular fever. Lymphocytes predominate in the tissue response to the tubercle bacillus, and are found surrounding malignant tumours. In conditions of lymphocytosis, hyperplasia of the red marrow occurs, and lymphocytes form the predominant element found there histologically.[1]

In *infectious mononucleosis* the total white count rises, *e.g.* to 12,000 per c.mm. and 50–70% of the circulating cells may be abnormal-looking mononuclear cells—probably atypical lymphocytes. *Monocytes* are increased in some virus diseases and in protozoal infections such as malaria.

Control of Leucopoiesis.[2]—Comparatively little is known about the factors responsible for the maturation of the white blood cells or those which regulate their discharge into the circulation. Some of the factors regulating *granulopoiesis* will be considered here. Sabin observed that following the

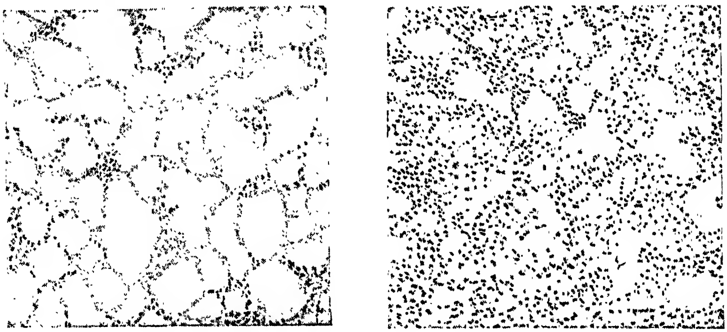
A  B

FIG. 124.—Effect of Nucleotide Injections on the Bone Marrow of the Normal Rabbit. (Doan, *J. Amer. med. Assoc.*, 1932, *99*.)

A.—Normal marrow from upper end of tibia removed at biopsy prior to injections
B.—Marrow from upper end of opposite tibia removed after 5 daily injections of nucleotide (0·35 g. each). Note the decrease in the fat spaces and the increase in cellularity which is due to hyperplasia of the myeloid cells.

appearance of showers of senile granulocytes in the circulating blood numerous young motile cells could be found. It was suggested that some product of granulocyte disintegration stimulates the turning-out of new cells from the marrow, the normal count thus being maintained. As nucleic acid is an important constituent of all nuclei, observations were made on the effects of injecting this substance and its derivatives into normal animals. Sodium nucleinate, pentose nucleotide (which is normally present in blood), adenylic or guanylic acid, adenine and guanine salts all induce a marked neutrophil leucocytosis and no toxic effects. The condition can be kept up for many weeks; no immature forms appear in the blood. The red marrow shows intense *myeloid hyperplasia* (Fig. 124), and new centres of white cell formation may appear in the spleen or kidney. *Pentose nucleotide* has been used in the treatment of agranulocytosis (p. 222); it is uncertain, however, whether any members of this group are concerned with the *normal* control of granulopoiesis.

[1] In the *leukæmias*, primitive white cells appear in the circulating blood.
[2] Doan, *J. Amer. med. Assoc.*, 1932, *99*, 194.

Inflammatory exudates in mammals and man contain protein constituents which influence white cell formation.[1] A pseudo-globulin which is a *leucocytosis-producing factor* has been isolated; when injected into animals it causes great extension and hyperplasia of the myeloid tissue in the marrow and a discharge of granulocytes into the circulation, leading to an increase up to fourfold in the circulating leucocytes, many of which are young cells. It is likely that this globulin plays a part in the normal response of the bone marrow to infections. Another agent, a polypeptide, has been isolated from inflammatory exudates which induces a transient *leucopenia* when injected into dogs. The way in which the leucocytes disappear is not clear, but it is assumed that they are "trapped" in various organs, *i.e.* the lungs, the liver,

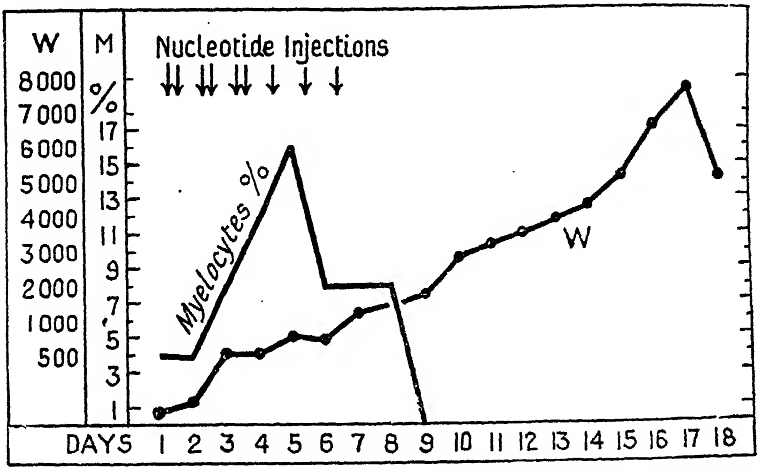


FIG. 125.—Treatment of Agranulocytosis with Nucleotide. (Doan, *J. Amer. med. Assoc.*, 1932, *99*.)

W = absolute numbers of granulocytes and their precursors (= myeloid cells) per c.mm. M = myelocytes shown as a percentage of total myeloid cells.

Note the initial increase in the myelocyte count which heralds the increase in granulocyte count.

and especially the spleen (perhaps accounting for the enlargement of the spleen in certain inflammatory states). It is probable that the white cell changes in the marrow and blood in inflammation depend largely on the nature of the products absorbed from the affected tissues (p. 223).

Agranulocytosis.[2]—In this disease the circulating granulocytes greatly decrease in number or almost completely disappear; the lymphocytes are unaffected. Agranulocytosis is often due to the toxic action of drugs, *e.g.*, sulphonamides, amidopyrin, arsphenamine, gold or thiouracil. The patient is usually gravely ill and shows severe throat infections; commonly there is *bacterial invasion of the blood* which may be the cause of death. Examination of the marrow proves that there is failure of granulocyte formation. In some cases (*aplastic type*) myeloid cells are not visible in the marrow, and no differentiation of the reticulum cells is taking place; this group is at present

[1] Menkin, *Dynamics of Inflammation*, 1940; *Amer. J. Path.*, 1943, *19*, 1021.

[2] Fettes and Whitby, *Lancet*, 1935, i, 205. Israëls and Wilkinson, *Quart. J. Med.*, 1937, *6*, 35.

not amenable to treatment. In others, myeloid development proceeds as far as the myeloblast or myelocyte stage, but *mature granulocytes are not formed* and are therefore not delivered into the blood (compare the imperfect maturation of the megaloblasts in pernicious anæmia (p. 196)). In some cases of the second group, the injection of nucleotide leads to an active response on the part of the marrow and finally to a restoration of the granulocyte count to normal. During the first few days of treatment immature cells (myeloblasts and myelocytes) appear temporarily in the blood stream (forming up to 15% of the granulocyte count); this is followed on about the fifth day by the entrance of young granulocytes and later of mature granulocytes into the blood and by an increase in the total white count (Fig. 125). The heralding of impending improvement by the transient appearance of immature white cells resembles the reticulocyte response seen in cases of pernicious anæmia treated with liver; this likewise precedes the rise of the red cell count. Nucleotides and related substances thus promote the maturation of the granulocytes in this disease, as well as stimulating the myeloid cells to proliferate more rapidly in normal subjects.[1] The associated infection is treated with a suitable drug, *e.g.* penicillin.

Inflammation.[2]—The cardinal features of inflammation are local heat, redness, swelling, and pain. The *rise of temperature* is due to increased blood flow from arteriolar dilatation; the *redness* is due to capillary dilatation; the *swelling* is due to excessive exudation and accumulation of protein-rich fluid from the more permeable vessels in the tissue-spaces (*inflammatory œdema.*) The vascular changes are due to products of tissue damage or of bacterial disintegration or toxins. The outflow of protein-rich plasma carries the protective immune bodies (associated with the γ-globulins) out of the blood to combat the organisms or their toxins in the tissues; among these are antibacterial and antitoxic substances, agglutinins (which clump the bacteria together) and opsonins (which prepare the bacteria for ingestion by phagocytes). At first, because of the arteriolar dilatation, the blood courses rapidly from the arteries to the veins. With progressive capillary dilatation and local hæmoconcentration the blood stream is increasingly slowed. The corpuscles at first remain in the centre of the capillary as an *axial stream*; but the leucocytes somehow become adherent to the capillary lining and are prevented from being swept on. They soon pass out through the seemingly intact capillary wall (*diapedesis*) and escape into the crevices of the tissues, where they wander about.

According to Menkin[3] all the phenomena of inflammation are due to specific agents which are released from the tissues, which have been injured by the bacterial toxins; he probably underestimates the part played *directly by bacterial toxins* on the blood vessels, blood-forming tissues, and the organs generally.

(i) Menkin believes that a diffusible chemical substance which attracts the leucocytes is liberated from the dead tissue cells and the dead bacteria; he has isolated from inflammatory exudates a polypeptide (*leucotaxine*) which exerts a chemical attractive influence (*chemotaxis*) on the granulocytes

[1] A marked decrease in the granulocyte count may result from excessive destruction of these cells in splenic overactivity ("hypersplenism," p. 229).

[2] Menkin, *Physiol. Rev.*, 1938, *18*, 366; *Arch. Path.*, 1943, *36*, 269.

[3] Menkin, *J. exp. Med.*, 1938, *67*, 129 *et seq.*

causing them to migrate out of the capillaries; it also increases capillary permeability. A similar substance is present in tryptic digests of serum; leucotaxine may therefore represent an intermediary product of tissue-protein hydrolysis.

(ii) Probably bacterial toxins have similar actions.

(iii) The bacteria in the tissues are dealt with by the leucocytes and the antibodies which have exuded from the plasma. The débris is liquefied and finally removed by tissue scavenger cells (monocytes, histiocytes).

(iv) As was explained on p. 221, Menkin claims that a *leucocytosis producing factor* passes from the site of infection into the blood stream to stimulate the myeloid tissue in the red bone marrow; a *leucopenic factor* is alleged to cause the "trapping" of circulating leucocytes. According to the nature of the infection, the proliferation of different kinds of white cells in the bone marrow is stimulated and different cells are "trapped," leading to a characteristic white cell picture in the circulating blood; similarly, different types of cells pass out of the blood vessels into the inflamed area.

(v) Two other agents described by Menkin may be released from damaged tissues at the site of infection: *necrosin*, which further damages the tissues locally and when it enters the general circulation damages organs elsewhere; *pyrexin*, which disturbs the temperature-regulating mechanism and induces fever (p. 480).

THE SPLEEN[1]

Structure of the Spleen.—The spleen is covered by a capsule which contains fibrous tissue and unstriped muscle. These plain muscle fibres contract at regular intervals (several times per minute) squeezing blood out of the spleen and thus driving it on towards the liver. The arteries pass in at the hilum surrounded by connective tissue, and soon divide repeatedly; the smaller branches become surrounded by a mantle of lymphoid tissue which swells out in places to constitute larger nodules, the *Malpighian corpuscles* (*bodies*). These are scattered fairly uniformly through the substance of the organ and usually have a vessel in close relation to them. The intervening tissue or *splenic pulp* consists of a framework of fibres derived from the fibrous trabeculæ which pass in from the capsule and accompany the incoming arteries.

The special cells found in the pulp are (Fig. 126):

(*a*) Large amœboid phagocytic cells which are part of the *macrophage* (reticulo-endothelial) system; (*b*) reticulum cells which show characteristic responses in various pathological states (p. 161); (*c*) a few giant cells.

Vascular Arrangements.—There is still considerable uncertainty about the vascular arrangements in the spleen; Knisely's views are as follows: after passing through the Malpighian body the small artery breaks up into a bunch of arterioles with well-developed muscular coats which constitute the *afferent sphincter*; the arterioles lead into a network of intercommunicating capillaries or *sinuses* which drain into collecting venules, and then into small veins. At the junction of sinus and venule there is a further aggregation of circular muscle fibres which constitute the *efferent sphincter*. Some

[1] Pearce, Krumbhaar, and Frazier, *Spleen and Anæmia*, Philadelphia, 1918. Barcroft, *Lancet*, 1925, i, 319. Krumbhaar, *Physiol. Rev.*, 1926, *6*, 160.

workers believe that there is free communication through *apertures* of varying size between the blood in the sinuses and venules with the interior of the pulp. Fig. 126 illustrates this view and shows the pulp packed with red cells, many of which have been taken up by macrophages ; it also shows the free passage of blood cells from pulp to venules and from venules to pulp. Knisely, from his studies of the living spleen, believes, on the contrary, that the endothelial lining of the splenic sinuses and venules is quite intact and that under normal conditions blood cells cannot pass out into the pulp. He describes the following remarkable variations in the blood flow through the smaller splenic vessels (Fig. 127) (but it must be emphasized that his findings still lack independent confirmation) :

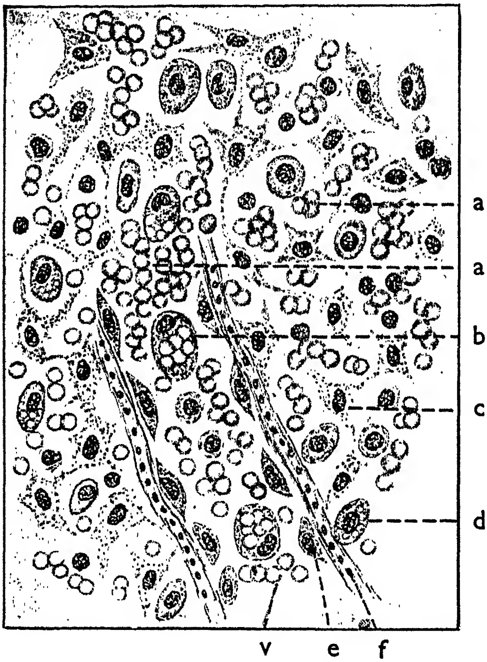


Fig. 126.—Structure of Splenic Pulp. (Sharpey-Schafer, *Essentials of Histology*.)

a,a =red cells in pulp and venule ; *b*, *d*=reticulo-endothelial cells (macrophages) in pulp and venule containing ingested red cells ; *c*=reticulum cell ; *v*=venous sinus ; *e*=endothelial lining ; *f*=network of fibres outside venule.

(i) There may be a rapid blood flow from afferent arterioles to venules through narrow capillaries.

(ii) *Filling Phase.*—The efferent sphincter contracts preventing the outflow from capillaries to venules. The capillaries become progressively distended with blood and can now be called sinuses. Their high internal pressure drives plasma out into the pulp, so that the sinus becomes a sack tightly packed with red corpuscles.

(iii) *Storage Phase.*—The afferent sphincter closes so that the sack of corpuscles is shut off from the circulation ; negligible numbers of red cells escape into the pulp.

(iv) *Emptying Phase.*—After minutes or hours the sphincters relax and the blood pressure drives the thick paste of cells into the venous circulation. Red cells only appear in the pulp and undergo phagocytosis there as a result of trauma to the spleen, or in dying animals.

The points of agreement in these conflicting descriptions are : (i) the spleen can *store red blood cells* at times (in the pulp or in the sinuses) and then discharge them into the general circulation ; the magnitude of this storage function varies with the species and is *small in man* ; (ii) during storage the red cells are either broken down by macrophages or, while separated from the plasma in the sinuses, they may be so modified that they are subsequently more liable to undergo disintegration. The main point of difference is, of

course, whether the red cells come in direct contact with the cellular elements of the pulp.

WEIGHT OF SPLEEN.—*Average* data from people who have died of accidents show that the spleen increases in weight progressively during the period of body growth to an adult level of about 150 g.; there is some decline during later life. The *individual* weights in such series show, however, a marked degree of scatter; most normal adult spleen weights lie between 50 and 250 g. with a few under 50 g., and a few at 300–350 g. The wide range of

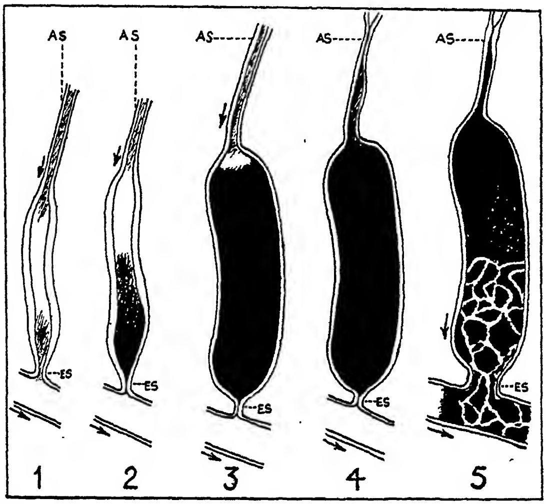


FIG. 127.—Changes in Blood Sinuses of Transilluminated exteriorized Spleen, showing Filling, Storage and Emptying Phases.

A.S. = afferent sphincter on afferent arterioles.
E.S. = efferent sphincter at junction of capillary (sinus) and venule.
1. Free flow from afferent arteriole through sinus to venule (*conduction* phase).
2. Efferent (distal) sphincter beginning to close; sinus filling.
3, 4. Efferent sphincter shut. Sinus greatly distended, plasma escapes, red cells retained. In 4, afferent (proximal) sphincter also closed = *storage* phase.
5. *Emptying* phase; efferent (distal) sphincter relaxes and the thick paste-like contents of the venous sinus are discharged into the venule. (Knisely, *Anat. Rec.*, 1936, *65*, 38.)

weight may in part be related to the varying amounts of blood present in the organ at death.

Functions of the Spleen.—1. BLOOD FORMATION.—(i) During the second half of fœtal life the spleen is actively engaged in forming red blood corpuscles (p. 160). In states of emergency, as in anæmia caused by destruction of the bone marrow this function may be resumed in post-natal life. (ii) The Malphighian corpuscles are characteristic nodules of lymphoid tissue, and by their proliferation add to the number of lymphocytes in the circulating blood (p. 218).

In the *leukæmias* the normal cells of the splenic pulp disappear, and are replaced by masses of lymphoid cells in lymphatic leukæmia, and by numerous myeloblasts or myelocytes in the myeloid type of leukæmia.

2. Blood Destruction.—The macrophages of the spleen are believed to destroy red cells (p. 187), lymphocytes, and blood platelets (p. 155). The spleen is undoubtedly a site of hæmoglobin breakdown and of bilirubin formation (p. 188). The spleen may make the red cells which are stored in its sinuses or in the pulp more spherocytic and so more liable to undergo hæmolysis (*infra*).

3. Reservoir of Red Corpuscles.—In some mammals (*e.g.* cat, dog) the spleen is an important reservoir of red blood cells, which are discharged into the general circulation in states of emergency associated with *anoxia*, thus increasing the oxygen-carrying power of the blood. Anoxia acts on the central nervous system, causing a discharge of impulses along the sympathetic nerves to the spleen and the adrenal medulla; under the combined nervous and hormonal stimulation the spleen contracts, by virtue either of the smooth muscle in the capsule and trabeculæ, or by the emptying of the filled blood sinuses. Anoxic conditions which produce these effects are: (i) diminished oxygen tension in the inspired air (*e.g.* high altitudes); (ii) carbon monoxide poisoning; (iii) hæmorrhage; (iv) severe muscular exercise (Fig. 128). Emptying of the spleen also results from mechanical stimulation, from raised external temperature which is associated with an increase in plasma volume and dilution of the cells of the blood, during œstrus (heat) and pregnancy (thus providing blood for the turgid generative organs), and during ether and chloroform anæsthesia.

Fig. 128.—Contraction of Spleen (Cat) after Exercise. (Barcroft, *Lancet*, 1925, i, 319.)

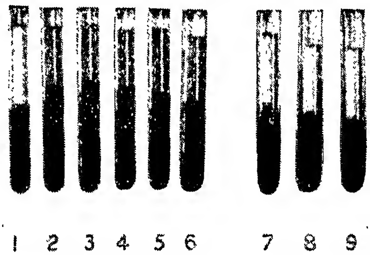


Fig. 129.—Increased Red Cell Concentration and Spontaneous Hæmolysis in Blood expelled from Resting Spleen. (Mellgren, *J. Physiol.*, 1939, *94*, 483.)

All the bloods were centrifuged *50 minutes after being drawn.*
Hæmatocrit tubes, 1 to 6: splenic venous blood taken at successive intervals during contraction of the spleen resulting from mechanical stimulation of the organ.
Tubes 7, 8, and 9 contain blood from inferior vena cava, externa jugular vein and carotid artery respectively; these serve as controls with respect to hæmatocrit value (*i.e.* corpuscular concentration) and plasma colour.
Note the higher red cell content in the splenic venous blood especially in tubes 2, 3, 4, 5, when the splenic contraction was maximal and the *blood previously stored in the organ was being expelled.* Spontaneous hæmolysis (causing darkening of the plasma) is present in the blood in tubes 2 to 6.

Stored blood expelled from the spleen has certain distinctive characteristics:[1]

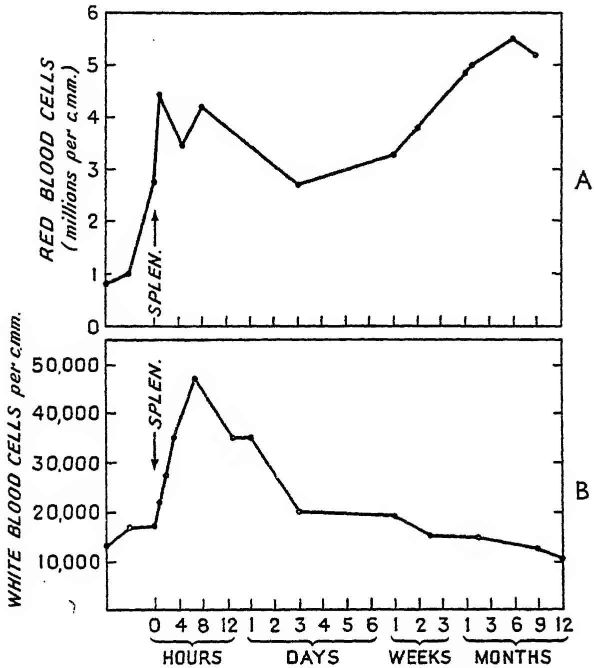
(i) Its *red cell content* is above that in the general circulation (Fig. 129).

(ii) The *sedimentation rate* is markedly slowed down (indicative of reduced rouleau formation).

(iii) The cells are more *spherocytic* (p. 170), and so undergo hæmolysis in hypotonic saline at concentrations as high as 0·65%.

[1] Watson and Paine, *Amer. J. med. Sci.*, 1943, *205*, 493.

(iv) On being allowed to stand *in vitro* the blood undergoes *spontaneous hæmolysis* to a significant extent (unlike blood from other vessels or organs): this point is well illustrated by Fig. 129 in which splenic vein blood was allowed to stand for 50 minutes after it was drawn before being centrifuged. These findings suggest that the storage phase may *prepare the red cells for hæmolysis* (in the spleen or elsewhere).



FIGS. 130 A and B.—Effects of Splenic Manipulation and Splenectomy in Case of Congenital Hæmolytic (Acholuric) Jaundice. (Modified from Sharpe *et al.*, *Arch. int. Med.*, 1939, *64*, 270.)

Upper Record: Changes in red cell count. Lower Record: Changes in white cell count. At arrow, Splen. = Splenectomy performed.

Note the enormous increase in circulating red cell and white cell counts produced by manipulation of the spleen at operation. The red cell count shows a temporary decline from the maximum and then rises as the beneficial effects of splenectomy develop. The white cell count slowly sinks to normal. [Other forms of anæmia do not show these reactions.]

(v) The red cells in the spleen may give staining reactions indicating that their *intracorpuscular* hæmoglobin has been broken releasing inorganic iron *in situ*.

The blood storage function of the spleen *in normal man* has not been convincingly demonstrated. Thus, though both injection of *adrenaline* and *muscular exercise* increase the red cell count in man, they produce comparable effects in splenectomized subjects and therefore the changes cannot be attributed to splenic contraction.

In cases of *congenital hæmolytic jaundice* (p. 229) the spleen can discharge

large numbers of red cells (and white cells) into the general circulation. When exposed at operation (for splenectomy) the organ is seen to be very large; during its mechanical manipulation it rapidly becomes smaller. This change is associated with a marked increase in red cell count, hæmoglobin concentration, and total white cell count of the peripheral blood. Typical results are shown in Fig. 130; the effects of splenectomy in this disease are considered on p. 229.

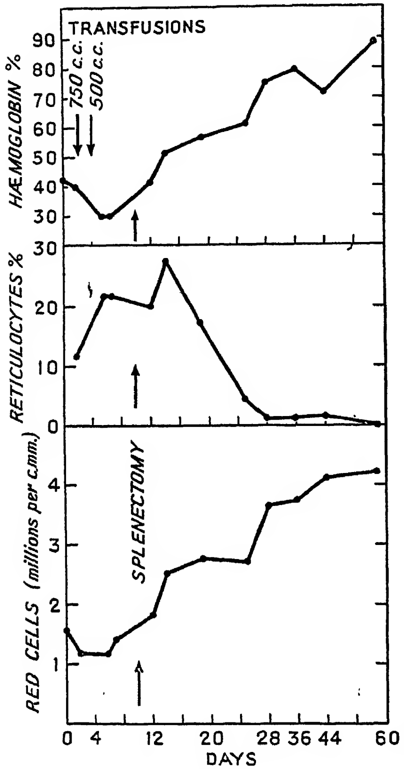


FIG. 131.—Effects of Splenectomy in Case of Grave Hæmolytic Anæmia with Circulating Hæmolysin. (Drawn from data by Dameshek and Schwartz, *New Engl. med. J.*, 1938, *218*, 76.)

Records from above downwards: hæmoglobin per cent. (100% = 15·5 g-%); reticulocytes per cent.; red cell count (millions per c.mm.).

On admission: grave anæmia (red cells 1·6 million, hæmoglobin 45%), evidence of hæmolysis (jaundice, raised serum bilirubin), and marked red marrow response (hyperplastic bone marrow, normoblasts predominating, reticulocytes 12–22%). Diagnosis of hæmolytic anæmia. Circulating hæmolysin subsequently demonstrated in serum. Condition rapidly deteriorated in spite of repeated transfusions (750 c.c. and 500 c.c. of blood injected at points indicated by arrows).

Splenectomy carried out when almost moribund. Rapid progressive improvement in state of blood and decline in reticulocytes. The removed spleen weighed 500 g., was congested, and contained infarcts; no increase in splenic macrophages.

4. RELATION TO HÆMOLYSIN FORMATION.—When red cells of a foreign species are repeatedly injected into an animal, specific hæmolysins are formed (p. 170) in the spleen and other organs probably by the reticulo-endothelial system. There are also rare clinical conditions in which the spleen forms hæmolysins that destroy the patient's *own* red cells. Fig. 131 shows the blood findings in such a case. The patient gave one month's history of progressive weakness, pallor, and jaundice. There was evidence of anæmia of the hæmolytic type, with increased serum bilirubin, and an active marrow response as shown by a high percentage of circulating reticulocytes and normoblastic hyperplasia in the bone marrow. The patient rapidly got worse in spite of repeated blood transfusions and liver therapy. When he was almost moribund splenectomy was performed; the removed organ weighed 500 g. The operation was followed by smooth recovery. A *hæmolysin* was demonstrated in the serum. In other cases the serum hæmolysin found was destructive to cells of all the blood groups including the patient's own cells, was peculiar in character, and was antagonized by incubation with normal human serum; a hæmolysin with similar properties was extracted from the spleen. Blood from this kind of case usually gives a positive *Coombs test* (p. 172).

The evidence just presented strengthens the suspicion that the spleen may normally be concerned with the

hæmolysis of red cells or with subjecting them to some treatment preparatory to actual hæmolysis.

5. Relation of Spleen to Purpura Hæmorrhagica.—See p. 159.

6. Defence Reactions.—There are many observations to show that the macrophages (reticulo-endothelial system) including those in the spleen, are actively concerned in the defence of the body against infection. Thus, if the macrophages generally are poisoned by injection of collargol,[1] the amount of *immune* bodies formed by the animal is diminished. Splenectomized animals cannot be immunized against tetanus toxin. After splenectomy animals succumb more readily to intercurrent infections than do the controls. The splenic cells unite readily with diphtheria toxin and retain it, and thus prevent it from exerting harmful effects on the body as a whole. Bacteria injected intravenously are taken up with great rapidity by the macrophages, which are also concerned with engulfing larger parasites, like those of kala-azar. It was pointed out (p. 218) that the macrophages in the germinal centres of lymph nodules destroy lymphocytes and so release their contained immune body (γ-globulin). In many subacute and chronic infections, *e.g.* in typhoid fever, the sinuses of the lymph nodes and the splenic pulp are packed with mononuclear cells which presumably deal with the invading organisms.

Congenital Hæmolytic (Acholuric) Jaundice.[2]—In this disease the cells are shorter and fatter than normal (*spherocytes*); as there is a reduced ratio of cell surface to cell volume the cells show increased fragility in hypotonic saline (p. 170 and Figs. 95, 96). If the cells from a patient are transfused into a normal person they have a *short survival time*, *e.g.* 14 days (normal about 100 days). No abnormal hæmolytic mechanism has been demonstrated in these patients. Because the cells are abnormal they are destroyed excessively rapidly in the patient's own body producing anæmia. From time to time and for no known reason erythropoiesis is temporarily suspended; the anæmia then becomes very severe as half the circulating red cells can be destroyed in 7 days. The serum bilirubin is raised in this disease and jaundice develops (*hæmolytic* jaundice, (p. 190)), but no bile pigment appears in the urine (hence the name "acholuric" (cf. p. 193)). The red marrow shows compensatory normoblastic hyperplasia and the reticulocyte count is high.

As already pointed out there is marked red cell storage in the spleen in this disease (p. 227 and Fig. 130). *Splenectomy* arrests the abnormal blood destruction; the serum bilirubin falls, the jaundice disappears, and the red cell count gradually rises to normal; but the spherocytosis and the increased cell fragility in hypotonic saline remain unaltered. It is not clear how splenectomy produces its beneficial effects.

Hypersplenism.—Clinically cases are encountered in which enlargement of the spleen is associated with suggestive evidence of overactivity of the organ, leading to *excessive destruction of blood platelets, granulocytes, and red-blood cells to a varying extent.* Some or all of the following changes may be found in the blood:

(i) A marked decrease in the *platelet* count (Fig. 132) sometimes associated with signs of *purpura* (p. 159).

[1] Collargol is a colloidal silver preparation.
[2] Dacie and Mollison, *Lancet*, 1943, ii, 550; Dacie, *Quart. J. Med.*, 1943, *12*, 101.

(ii) Decrease or disappearance of the neutrophil granulocytes (granulopenia, *neutropenia*), in spite of active myeloid proliferation and maturation in the bone marrow.

(iii) *Anæmia* due to excessive blood destruction (with raised serum bilirubin and jaundice, and sometimes with evidence of marked erythrophagocytosis in the spleen).

These patients are cured by *splenectomy*. Immediately after the operation there may be a marked overswing of the platelet count (*e.g.* to 2·5 million per c.mm.), and of the granulocyte count (*e.g.* to 20,000 per c.mm.); these reactions probably indicate that the hypertrophied bone marrow goes on manufacturing these cells in large numbers, and as they are no longer being

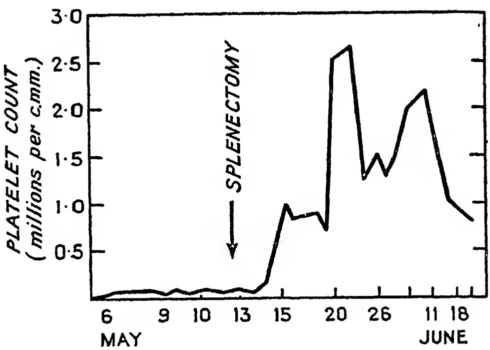


FIG. 132.—Effect of Splenectomy in Thrombocytopenic Purpura attributed to Hypersplenism. (Modified from Wiseman and Doan, *Ann. int. Med.*, 1942, *16*, 1100.)

Girl aged 12 suffering from purpuric hæmorrhages. There was moderate secondary anæmia, mild neutropenia, and *almost complete disappearance of blood platelets.* Following splenectomy the platelets rose rapidly to over 2·5 millions/c.mm. There was also a temporary granulocytosis.

excessively destroyed, their count in the blood temporarily rises to a high level; ultimately a normal state of the blood is attained.

Probably no hard and fast distinction should be drawn between this group labelled hypersplenism, cases of splenic anæmia with circulating hæmolysin, and cases of purpura hæmorrhagica (with low platelet count) that benefit from splenectomy. Our knowledge of splenic function in health or disease is too uncertain to allow dogmatic opinions on the subject.

Results of Splenectomy.—The main results of splenectomy in *normal animals* are:[1]

(i) An *anæmia* of the hypochromic type, mild or moderate in character, which usually reaches its severest stage after 1½ months and is followed by slow and steady recovery during the succeeding three or four months (Fig. 133, A). The red count rarely falls below 3,000,000 per c.mm. or the hæmoglobin below 55%. These results are quite unexpected in view of the relationship of the spleen to red cell destruction. Unfortunately there are few data on the subject in normal human beings. Krumbhaar believes that

[1] Singer and Weisz, *Amer. J. med. Sci.*, 1945, *210*, 301.

the spleen gives rise to some substance which normally *stimulates the marrow* to activity and that the anæmia is due to decreased marrow activity. The anæmia is later compensated for by hyperplasia of the red bone marrow.

(ii) The red cells are *thinner* and consequently are more resistant to the action of hypotonic saline (Fig. 96); their life span is *not* prolonged.

(iii) In animals in which the spleen is an important reservoir of red cells, there is diminished resistance to severe anoxia from any cause (*e.g.* hæmorrhage, CO poisoning).

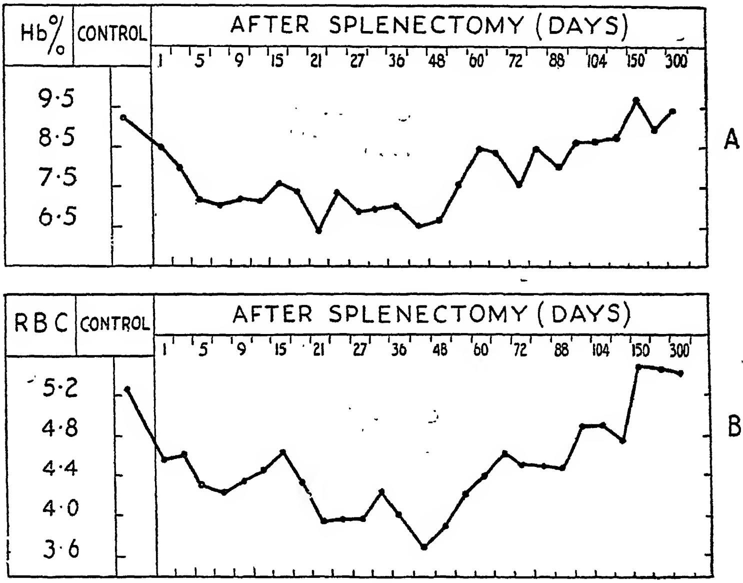


FIG. 133.—Changes in Hæmoglobin Values and Red Cell Count after Splenectomy in Normal Dogs. (Krumbhaar *et al.*, *Spleen and Anæmia.*)

Hb% : hæmoglobin concentration g. per 100 c.c. (A)
RBC : red cell count in millions per c.mm. (B

(iv) Blood destruction and bilirubin formation continue to an undiminished extent (p. 188) from compensatory activity of macrophages elsewhere.

(v) In some animals resistance to certain forms of infection is decreased.

(vi) There is an unexplained increase in the total number of *leucocytes*, perhaps to 20,000–40,000 per c.mm.; this is almost entirely due to the increase in the neutrophil count, which returns to normal very slowly (Fig. 133, B).

SPLENECTOMY IN MAN.[1]—This operation is beneficial in man in congenital hæmolytic jaundice (p. 229), hypersplenism (p. 230), splenic anæmia with circulating hæmolysin (p. 228), and some cases of purpura hæmorrhagica (p. 159).

[1] Learmonth, *Brit. med. J.*, 1951, ii, 67.

NOTE ON TERMINOLOGY OF ABNORMAL BLOOD CONDITIONS IN INFANTS AND CHILDREN.

The terminology of these conditions is liable to give rise to confusion; they are gathered together here and commented on briefly in the hope of producing some degree of clarity.

Haemolytic Disease (p. 180); a condition due to the hæmolytic action of the Rhesus antibody on the fœtus and new-born infant. Some of the manifestations of this disease are named as though they were independent clinical syndromes (which they are not). These names include: icterus gravis neonatorum; congenital anæmia of the new-born; erythroblastosis fœtalis (erythroblastæmia); hydrops fœtalis; kernicterus.

Icterus Neonatorum (Jaundice of New-born) (p. 192): probably due to functional immaturity of the liver.

Hæmorrhagic States in Infants (p. 153): possibly due to Vitamin-*K* deficiency.

Hæmophilia (p. 150).

Purpura (p. 157).

Congenital Hæmolytic Jaundice or Icterus (Congenital Acholuric Jaundice; Congenital Hæmolytic Anæmia) (p. 229): a congenital abnormality of structure of the red cells.

III

THE HEART AND CIRCULATION

STRUCTURE AND PROPERTIES OF HEART MUSCLE [1]

Cardiac Muscle.—The muscle fibres of the heart are cylindrical in shape, have a nucleus situated centrally, and branch, uniting by their branches with adjacent fibres (Fig. 134). They are longitudinally striated as is all

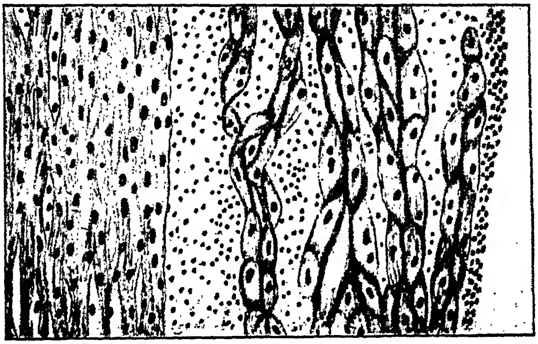


FIG. 134.—Structure of Heart. (Lewis, *Mechanism and Graphic Registration of Heart Beat.*)

Section through sheep's heart.
C. = Muscle fibres of ventricle.
P.F. = Purkinje fibres (separated by connective tissue.)
On the right-hand surface is the endothelial lining of the ventricle (E.).
Note the large size of the Purkinje fibres.

muscular tissue, and in addition they show some *transverse* striation. As no definite cell membranes are present between the individual fibres, heart muscle must be regarded as a syncytium. For all practical purposes there is functional continuity throughout the muscle tissue of the auricles and of the ventricles respectively.

Special Junctional Tissues of the Heart.[2]—This term is employed to describe certain tissues in the heart which are concerned with the initiation and propagation of the heart beat. They include the *sino-auricular node* and the junctional tissues—the latter consisting of the *auriculo-ventricular node*; the auriculo-ventricular bundle, or *bundle of His*; the right and left *divisions* of the bundle, and their arborizations under the endocardium

[1] " Consult, peruse and study " Lewis, *Mechanism and Graphic Registration of Heart Beat*, 3rd edn., London, 1925.

[2] Nonidez, *Amer. Heart. J.*, 1943, *26*, 577.

and the *terminal fibres* which penetrate the ventricular substance (Fig. 135). The bundle of His and all its ramifications may be called *Purkinje tissue* (or *fibres*).

(i) THE SINO-AURICULAR NODE is situated at the junction of the superior vena cava and the free border of the right auricular appendix, and extends down along the sulcus terminalis for a distance of 2 cm. It is 2 mm. in width and has a rich capillary blood supply. It consists essentially of thin, elongated muscle fibres (about one-third the size of heart-muscle fibres), fusiform in shape and longitudinally striated, which interlace with one another in a plexiform manner. These fibres normally *initiate the heart beat*; for this reason the sino-auricular node is called the *pacemaker*. Nerve cells and fibres forming the excitor relay of the vagus nerve (Fig. 135) and excitor (postganglionic) fibres of the sympathetic are also present (p. 268).

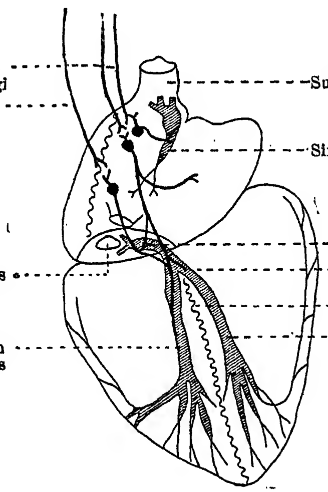


FIG. 135.—Diagram to show Arrangement of Special Tissues of Heart. (After Wiggers.)

(ii) THE AURICULO-VENTRICULAR NODE is situated at the posterior and right border of the inter-auricular septum near the mouth of the coronary sinus. Auricular muscle fibres from the region of the coronary sinus collect fanwise, interlace, and unite with the auriculo-ventricular node. In structure this node is identical with the sino-auricular node.[1]

(iii) THE BUNDLE OF HIS runs upwards to the posterior margin of the membranous part of the interventricular septum and then forwards below it, ensheathed and isolated in a canal. At the anterior part of the membranous septum, in front of the attachment of the septal cusp of the tricuspid valve

[1] The sino-auricular node is a right-sided structure developing from tissue which lies at the entrance of the primitive right great vein (which later becomes the superior vena cava). This explains the position of the node and its supply by the *right* vagus nerve. The auriculo-ventricular node is a left-sided structure, supplied therefore by the *left* vagus nerve, and developed from tissue in the vicinity of the entrance of the left great veins which becomes the coronary sinus. The anatomical position of the two nodes is thus also accounted for (Keith).

to the A.V. ring, the bundle forks. The *left* division pierces the membrane and then lies on the upper border of the muscular septum to enter the subendocardial space of the left ventricle beneath the union of the anterior and right posterior cusps of the aortic valve. This relation to the aortic valve is important, and explains how this branch of the bundle may be implicated in aortic incompetence. The *right division* passes down the right side of the septum and is mainly transmitted in the moderator band. Both branches are continued as an arborization of fibres lying under the endocardium of both ventricles from which terminal fibres penetrate the ventricular wall.

Purkinje tissue (bundle of His and its branches) differs histologically from cardiac muscle. In man the Purkinje fibres are somewhat larger (range 10–46 μ, mean 16 μ); the cell outlines are indistinct; the central cytoplasm is granular and contains several nuclei; the peripheral cytoplasm contains myofibrillæ but these are separated by more sarcoplasm; the glycogen content is greater. Intermediate forms between typical Purkinje fibres and cardiac muscle fibres are numerous. The bundle of His contains nerve fibres and many blood vessels.

Properties of Heart Muscle.—(1) Excitability and Contractility.—Heart muscle is *excitable*, *i.e.* it responds to external stimuli by *contracting*. The "all-or-none" law (p. 489) applies to the heart: if the external stimulus is too weak, no response is obtained; if the stimulus is adequate, the heart responds to the best of its ability. In this connection the auricles or ventricles behave as a single unit, so that an adequate stimulus normally produces a full contraction of auricles or ventricles. The force of the contraction obtained depends on the state in which the muscle fibres find themselves. The force varies with (i) the *initial length* of the fibres (p. 275); (ii) the duration of the previous *diastolic pause*; and (iii) the *nutrition* and *oxygen supply* (p. 236).

(2) Refractory Period.—(i) Throughout the period of contraction, heart muscle is *absolutely refractory* and does not respond at all to external stimuli. Summation effects are impossible, because one contraction must be completed and recovered from before a fresh one can be set up; as a consequence, heart muscle cannot develop a tetanus. The absolute refractory period of auricular muscle for heart rates up to 100 per minute is about 0·2 second.

(ii) Shortly after the contraction is over, the muscle is *relatively refractory*. Minimal stimuli fail to produce a response, but very strong stimuli are effective. The force of contraction during this period is also subnormal. Finally, full recovery occurs.[1]

(3) Conductivity is a property of all heart muscle, but is especially developed in the bundle of His and its branches (the Purkinje tissue). Conduction in Purkinje tissue is at the rate of 4 metres per second, in the auricular wall 1 metre per second, and in the ventricular wall 0·4 metre per second. In the A.V. node conduction is very slow—about 0·2 metre per second. (Large nerve fibres in mammals conduct at a rate of about 100 metres per second.) The excitation process is normally transmitted from one point in the heart to the adjacent one with no diminution of intensity, as in

[1] The refractory period is longest in the A.V. node, intermediate in the ventricles, and shortest in the auricles. Certain drugs, like digitalis and quinidine, prolong the absolute refractory period; stimulation of the vagus shortens the duration of systole, and so *diminishes* the refractory period.

nerve. When the Purkinje tissue is damaged, it conducts more slowly and with diminished intensity like narcotized nerve (p. 489).

(4) Rhythmicity.—The heart has the power of initiating its own impulse, without recourse to external agencies. Normally this function is carried out by the S.A. node, but under exceptional circumstances the heart beat may be initiated for long periods by the A.V. node or other parts of the junctional tissues (cf. p. 286).

Nutrition of the Heart.—This may be studied by perfusing the isolated heart through its coronary vessels. The heart is removed from the body, a cannula inserted into the aorta pointing towards the heart, and connected with a pressure bottle. The nutrient fluid under pressure is forced into the cannula, closes the semilunar valves, perfuses the coronary arteries, and thus reaches the heart muscle; it returns to the right auricle via the coronary sinus. The heart is attached to a recording lever and is soon found to contract rhythmically (note that the ventricles are empty). The rôle of the following factors can be demonstrated:

(1) Temperature.—Cooling the perfusing fluid slows the heart and may finally stop it; warming the fluid quickens the heart. These effects are due to alterations in the metabolic rate of the S.A. node.

(2) Pressure.—If the pressure of the perfusion fluid is lowered, the force of contraction is diminished (cf. p. 237).

(3) Oxygen Supply.—If this is inadequate, the beat becomes feeble and irregular and finally stops.

(4) Constituents of the Perfusing Fluid.—The inorganic constituents are more important than the organic. For the mammalian heart, fluid with the following salt concentration is best: NaCl, 0·9 g., $CaCl_2$, 0·024 g., KCl, 0·042 g., $NaHCO_3$, 0·01–0·03 g., distilled water, 100 c.c.

The part played by each constituent is as follows:

Na^+.—The NaCl is mainly responsible for giving the fluid its proper osmotic pressure so that it may be in equilibrium with the tissues. Na^+ ions must be present in the interstitial fluid to maintain the excitability of heart muscle (and skeletal muscle and nerve), its contractility and its rhythmic activity.

Ca^{++}.—If Ca^{++} ions are added to the perfusing fluid, the force of the heart is increased. If Ca^{++} is present in excess, the heart contracts well, but relaxes progressively less; the beats become progressively smaller in extent, and finally the heart stops in a condition of extreme contraction (*calcium rigor*).

K^+.—Excess K^+ causes increasing relaxation of the heart, and finally brings about arrest in complete diastole.

H^+ ion Concentration.—The fluid must be slightly alkaline. Increased H^+ ion concentration causes relaxation of the heart and feebleness of the beat; excessive alkalinity produces the reverse effect. Excessive artificial respiration in an intact animal results in washing out of CO_2 from the blood and consequent alkalæmia; the heart contracts forcibly but does not relax during diastole to receive the venous return, and so the output and blood pressure fall.

Metabolism of Heart Muscle.[1]—Cardiac muscle resembles skeleta. muscle in its chemical composition and in its metabolic processes (cf. p. 4216)

[1] Clark *et al.*, *Metabolism of Frog's Heart*, London, 1938. Evans, *Edin. med. J.*, 1939, *46*, 733.

Certain differences are noteworthy, the most important being that cardiac muscle cannot continue to beat after its O_2 supply has been used up, *i.e.* it cannot incur an oxygen debt (p. 439). The effects of cutting off the coronary blood flow are, therefore, grave (p. 238). Heart muscle can take up *lactate* as well as glucose from the blood.

The *oxygen consumption* of the heart varies with the amount of work it has to perform; it thus increases when the peripheral resistance, heart rate or output per beat rises (the other factors in each case remaining constant). The optimum mechanical efficiency of the heart is about 30% (cf. p. 431), and is observed when the cardiac output is high.

In the *intact* animal increased peripheral resistance is, however, often associated with reflex cardiac slowing; the work of the heart may then increase without increased oxygen usage, *i.e.* its efficiency increases under these circumstances. Efficiency also increases with rising output per minute up to a certain point; if the heart is overworked, however, it dilates and its efficiency falls off again.

Coronary Circulation.[1]—*Nerve Supply.*—The coronary vessels receive a rich innervation from both the vagus and the sympathetic. The vagus conveys *constrictor* fibres to the coronary arteries. Section of the vagi increases, and stimulation of the peripheral end of the vagus diminishes the coronary flow; in the latter case the same result is obtained even if the heart is artificially maintained at its previous rate by means of electrical stimulation. Stimulation of the sympathetic *dilates* the coronary arteries; cerebral anæmia and asphyxia produce coronary dilatation, in part by stimulation of the sympathetic supply to the coronary vessels.

The blood supply to the heart is adjusted to its varying needs, the main factors responsible being *oxygen tension, CO_2 tension* (and *H^+ ion concentration*), *mean blood pressure, nervous reflexes, temperature, adrenaline,* and certain *non-acid products of activity.*

(1) *Blood Composition.*—Anoxia greatly increases the coronary flow; a fall in the oxygen saturation of the arterial blood to 80% has little effect, but a fall to 50% may increase the flow four- to fivefold. *Excess CO_2* has a much smaller effect. A change in the CO_2 content of the inspired air from 3 to 7% only increases the coronary flow by 50%. Adrenaline, as might be expected, dilates the coronary arteries in most species (p. 725).

(2) *Influence of Arterial Blood Pressure.*—The mean aortic blood pressure[2] is one of the main factors which control the coronary circulation. Thus in the denervated heart-lung preparation (dog) (p. 276) an increase in mean blood pressure from 50 to 130 mm. Hg may increase the coronary flow from 20 to 250 c.c. per minute. *A low blood pressure* from any cause similarly leads to an impaired blood supply through the coronaries and defective nutrition of the heart, which further depresses the state of the circulation.

(3) *Influence of Cardiac Output.*—If the nerve supply of the heart is intact it is found that the *coronary flow increases as the cardiac output increases.* Anrep attributes this response to *reflex inhibition* of the vagal vasoconstrictor fibres to the coronary arteries.

(4) *Coronary flow during the Various Phases of the Cardiac Cycle.*[3]—This

[1] Gregg, *Physiol. Rev.*, 1946, *26*, 29; *Coronary Circulation in Health and Disease*, 1950.
[2] By "mean blood pressure" is meant the mean of systolic and diastolic pressure.
[3] Green, Gregg, and Wiggers, *Amer. J. Physiol.*, 1935, *112*, 627.

depends on the resultant of two main factors: the *aortic blood pressure* and the *state of the coronary blood vessels.* The main coronary *inflow* occurs during the period of ventricular *diastole,* while the main *outflow* occurs during ventricular systole. Closer analysis shows that three phases can be recognized (Fig. 136): During the isometric contraction phase the aortic blood pressure is minimal and the intramuscular branches of the coronary vessels are firmly compressed by the contracting wall of the ventricle. The coronary inflow is consequently sharply reduced (*a*) (but the *outflow* is correspondingly increased). During the ejection phase the coronary inflow follows the rise and fall of aortic blood pressure (*b*). During isometric relaxation the coronary inflow mounts steeply because the aortic pressure is high and the ventricular wall is now relaxed (*c*). The inflow, however, goes on rising to about the middle of the diastole (for no very clear reason) although the aortic pressure is falling. The inflow declines during the rest of diastole with the decrease in diastolic pressure.[1]

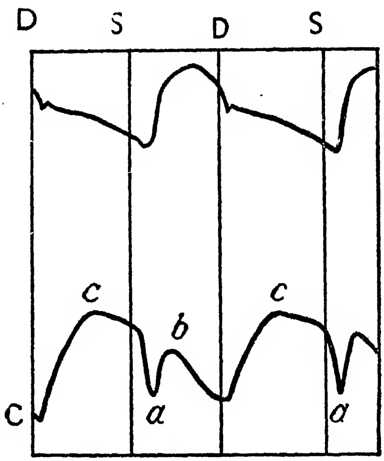


FIG. 136.—Changes in Coronary Inflow during Phases of Cardiac Cycle.

Upper record A: Aortic pressure tracing. Lower record C: Coronary inflow. Vertical lines S, D, mark onset respectively of ventricular systole and diastole. *a, b, c*=phases of coronary flow. (Green, Gregg, and Wiggers, *Amer. J. Physiol.*, 1935, *112*.)

(5) *Coronary Flow in Exercise in Man.*—This can be calculated with a fair degree of accuracy and turns out to be about *one litre per minute* in the most violent exercise.[2] This immense coronary flow may be attributed in the light of the results discussed above to the following factors: local rise of CO_2 tension and of H^+ ion concentration and the liberation possibly of non-acid dilator metabolites (cf. p. 316); low oxygen tension in the heart; secretion of adrenaline; reflex inhibition of the vagal vasoconstrictor fibres; raised systemic blood pressure.

RESULTS OF CORONARY OCCLUSION.[3]—Experimentally it is found that if both main branches of the left coronary artery are tied, death usually results

[1] Compare p. 432 for the blood flow changes during contraction in skeletal muscle.

[2] The maximum output of the heart per minute in exercise is about 40 litres. Let us assume the mean blood pressure to be 100 mm. Hg. The *work* of the left ventricle (output × resistance) can then be calculated and is found to be equivalent to 128 g.-calories (the *small calorie* is here referred to). If the mechanical efficiency of the heart is taken at a maximum of 33%, the total energy liberated by the chemical changes is $128 \times 3 = 384$ g.-calories per minute. If the work of the right ventricle is one-fourth that of the left (owing to the low pulmonary pressure), the total energy output of the heart is $384 + 96 = 480$ g.-calories per minute. 96 c.c. O_2 are needed to liberate this energy (1 c.c. $O_2 = 5$ g. calories). If 12 c.c. O_2 are obtained from each 100 c.c. of coronary blood, the flow through the coronary arteries must be at least 800 c.c. *per minute* (Hill).

[3] The detailed distribution of the coronary arteries is described by Gross (*Blood Supply of Heart,* 1921, New York); the anterior descending branch of the left coronary is particularly important because it is occluded more often than all the other branches added together. This branch supplies portions of *both* ventricles, *i.e.* the apex

immediately from ventricular fibrillation (p. 294). If the circumflex or anterior descending branch alone is ligated, no serious circulatory changes may be noted. If the right coronary artery is ligated, about half the experimental animals may survive for longer or shorter periods.

Clinically coronary occlusion may prove fatal instantaneously or within a few minutes, hours, or days, or death may occur later from heart failure, or good recovery may take place. The condition is characterized by *pain* which may be severe and prolonged, persisting for hours or days; it may be relatively unaffected by morphine. The symptoms of coronary occlusion are fully considered on p. 750, to which reference should be made. The referred pain in this condition has a very wide distribution, and may be accompanied by tenderness and rigidity of the abdominal wall, nausea, and vomiting. There are marked signs of *shock*: ashen-grey hue, sweating, and a very feeble pulse. The *blood pressure* falls markedly, sometimes to an extent which interferes with the secretion of urine. The electrocardiographic changes are considered on p. 251. The usual signs of heart failure may develop.[1]

ORIGIN AND SPREAD OF THE CARDIAC IMPULSE. ELECTROCARDIOGRAPHY[2]

Electrical Changes in Heart.—Electrical methods have proved of great value in the study of the origin and propagation of the excitation process in the heart and in the investigation of clinical disorders of the heart's action. The special instrument employed for these purposes is the *electrocardiograph.*

When nerve or skeletal muscle is stimulated, the active region becomes electrically negative with respect to a resting region, resulting in a characteristic deflection (p. 484); the same is true of heart muscle. The duration of the electrical disturbance varies with the tissue; in nerve fibre it is one or a few milliseconds (p. 490); in skeletal muscle it is a little longer but coincides approximately with the latent period, the electrical change being completed by the time that muscle tension begins to develop (p. 502). In cardiac muscle, however, the electrical disturbance persists *throughout the period of systole* and some electrical change may be present even in diastole;

---

of the left ventricle, the left anterior third of the right ventricle, and part of the interventricular septum. There is a rich anastomosis between the branches of the same artery and between the right and left coronaries in their capillary and precapillary distribution. The vascular needs of heart muscle are so enormous, however, that if a sufficiently large branch is blocked, death of the corresponding region of the heart (infarction) may take place. If a coronary vessel is slowly narrowed an adequate collateral circulation is gradually established so that the vessel may finally be completely closed without harm. According to Gross, anastomoses throughout the coronary circulation become more effective and more extensive with advancing years, so that, other factors being equal, the heart of a man at 60 is better prepared to withstand the effects of coronary occlusion than the heart of a man at 20. The results of coronary occlusion therefore depend on the size and location of the vessel, the rate of formation of the thrombus, the age, and the state of the general circulation.

[1] Ratnoff, *Medicine*, 1946, *25*, 285.

[2] Lewis, *Mechanism and Graphic Registration of Heart Beat*, 3rd edn., London, 1925. Goldberger, *Unipolar Lead Electrocardiography*, Phila., 2nd edn., 1949. Nahum and Chernoff, in Fulton, *Textbook of Physiology*, 16th edn., 1949. Hill, *Lancet*, 1950, i, 985, 1027.

stage of recovery with the T waves of the electrocardiogram of the entire ventricle when *indirect* leads are used. No one knows exactly how the *monophasic* single fibre potential is transformed into the *complex* waves of the ventricular cardiogram.[1] The following points should be noted: the arrangement of the ventricular fibres is extremely complicated; the exact

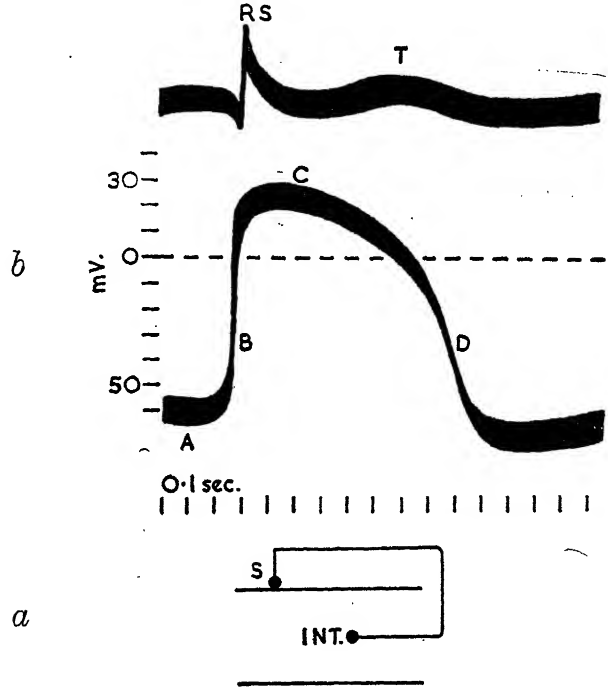


FIG. 137.—Action Potential of Single Heart Muscle Fibre (Frog Ventricle). (After Woodbury, *et al.*, *Circulation*, 1950, *1*, 264.)

(*a*) Disposition of electrodes.
S = surface electrode. Int. = electrode in interior of fibre.
(*b*) Below: action potential of single fibre. Ordinate: millivolts (mV). Zero base line: potential of surface of *resting* fibre. Downward deflection: relative negativity of interior of fibre. Upward deflection: relative negativity of surface.
A, fibre at rest. B, stage of excitation (invasion, "depolarization"). C, stage of possession. D, stage of recovery (retreat, "repolarization").
Above: electrocardiogram of frog ventricle (one cycle) using indirect leads.

course of the invasion process through the ventricular wall is not known with certainty, especially in man; the duration of the phase of possession in any region is unknown; the exact sequence and rate of repolarization are obscure. It is, therefore, impossible to describe the moment to moment changes in the pattern of current flow in the heart except in fairly general terms. To this extent the basic analysis of electrocardiographic curves is imperfect even in the case of the normal heart.

[1] Note also that the voltages recorded from the whole heart by indirect leads are small compared with the directly recorded single fibre potential.

The currents which are generated in the heart are conducted throughout the body fluids; from the electrical standpoint the heart can be regarded as being suspended in a bag of saline which acts as a volume conductor. If the chest is opened, or if the medium round the heart is altered in other ways (*e.g.* by pericardial effusion, pneumothorax, or the presence of emphysema) the pattern of current spread through the body fluids is altered and the electrocardiogram recorded by indirect leads is correspondingly modified.

Electrocardiographic Leads.—Two kinds of indirect leads are employed: *unipolar* and *bipolar*.

1. **Unipolar Leads.**—One electrode, the exploring electrode, is placed on an area of the body surface. The other, or *indifferent* electrode, is kept at approximately zero potential by connecting electrodes, placed respectively on the right arm, left arm, and left leg, to a central terminal through a 5000 ohm resistance. (The currents from the three limbs neutralize one another.) The indifferent electrode undergoes no significant change during the cardiac cycle. When a unipolar lead is used the electrocardiogram records the potential changes which affect the exploring electrode only.

The following unipolar leads are used; they are labelled V, followed by a letter or by a letter and number describing the position of the exploring electrode.

Name of Unipolar Lead.	Position of Exploring Electrode.
VR . .	Right arm.
VL . .	Left arm.
VF . .	Left leg (F=foot).
VC . .	Chest.
	Six chest positions are described:
VC 1 . .	4th intercostal space to right of sternum.
VC 2 . .	4th intercostal space to left of sternum.
VC 3 . .	Midway between left sternal border and mid-clavicular line on a line joining positions 2 and 4.
VC 4 . .	5th intercostal space in midclavicular line.
VC 5 . .	5th intercostal space in left anterior axillary line.
VC 6 . .	5th intercostal space in left midaxillary line.
VC 7 . .	5th intercostal space in left postaxillary line.

The chest leads VC 1–7 are sometimes called V 1–7.

The wiring employed is by convention such that when the exploring electrode is negative relative to the central zero terminal the record is deflected downwards; when the exploring electrode is positive the record is deflected upwards.

(1) UNIPOLAR LIMB LEADS.[1]—Each unipolar lead can be regarded as being projected as a solid cone on to the surface of the heart. The cardiac area subtended by the cone and which thus "faces" the electrode is called the *proximal* zone; most of the rest of the heart, which "faces away" from the electrode, is called the *distal* zone. There is a narrow *intermediate* zone between

[1] Nahum *et al.*, *Amer. J. Physiol.*, 1948, *153*, 529, 540, 547; *154*, 369.

the above mentioned zones (cf. Fig. 142). There is evidence that negativity in the proximal zone causes a downward movement of the record. Changes in the intermediate zone have little effect. The record, therefore, moves up or down when negativity predominates in the distal or proximal zone respectively. This view of the origin of the unipolar lead electrocardiogram is illustrated in Fig. 138. When excitation begins in the proximal zone, the negativity there produces a downward deflection. As excitation spreads to involve the distal zone the record begins to move upwards again towards the base line (*i.e.* the isoelectric line). As recovery associated with positivity develops in the proximal zone there is a deflection upwards above the base line. As the distal zone also recovers and becomes positive the record returns again to base line. It is believed that only electrical changes at the pericardial surface affect the indirect leads.

Fig. 142 shows the approximate distribution of the proximal and distal zones in the human heart when the unipolar limb leads VR, VL, and VF are employed.

(2) UNIPOLAR CHEST LEADS.[1]—Fig. 143 shows the position of the proximal zone in man in the case of each of the six unipolar chest leads. The proximal zone is much smaller than in the case of the limb leads and is roughly the portion of the heart which lies under the exploring electrode; there is overlap of the proximal zone of one position with those on either side of it. As the position of the exploring electrode recedes away from the ventricular surface (in chest positions 5, 6) the proximal zone becomes larger. In all the chest positions of the exploring electrode the "neutral" intermediate zone too is large. Negativity at the chest electrode is recorded as a *downward* deflection.

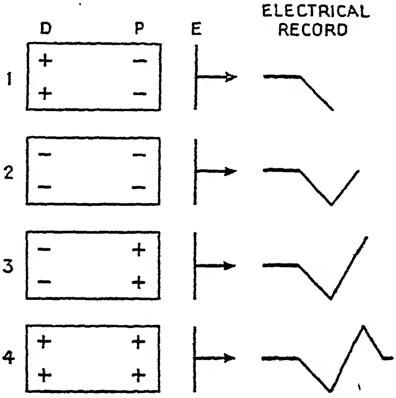


FIG. 138.—Mode of Production of Unipolar Lead Electrocardiogram.

D, P: distal and proximal zones of muscle mass.
E: electrode of unipolar lead.
1. Excitation associated with negativity begins in proximal zone: deflection downwards.
2. Owing to spread of excitation, negativity also develops in distal zone: record returns to base-line.
3. Recovery begins in proximal zone: deflection upwards.
4. Recovery extends to distal zone: record returns to base line.

2. **Bipolar Leads.**—(1) CLASSICAL LIMB LEADS.—These are the leads which have been most extensively used in man; a vast amount of information has consequently accumulated correlating the electrical with the clinical and post-mortem findings. These leads were also used in the earlier fundamental studies of the origin and spread of the cardiac impulse and in the analysis of experimentally induced abnormalities of the heart's action. The classical limb leads are:

Lead I: from right arm and left arm.
Lead II: from right arm and left leg.
Lead III: from left arm and left leg.

[1] Nahum and Hoff, *Amer. J. Physiol.*, 1948, *155*, 215.

By convention, negativity at the cranial or at the right-hand electrode is recorded as an *upward* deflection. Representative records are shown in Fig. 139.

These limbs are still the most useful for the majority of clinical purposes; they are, however, difficult to account for in terms of current flow in the heart. *Each* of the two electrodes in the bipolar leads is affected by the related

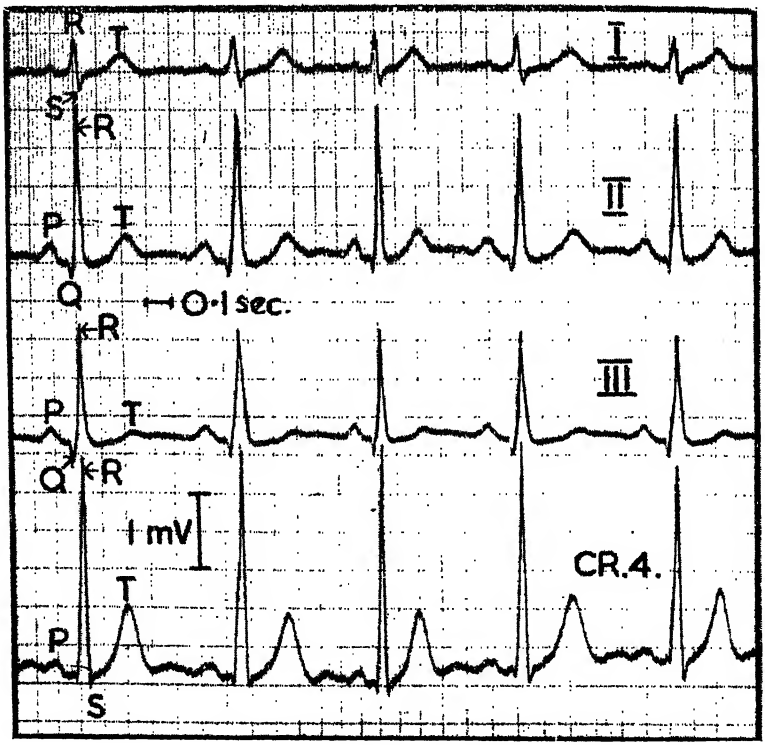


Fig. 139.—Simultaneous Electrocardiograms in Normal Person. Leads I, II, III and Chest Lead[1] (CR 4) (Cardiographic Department, Middlesex Hospital.)

proximal and distal zones. The deflection recorded thus represents the *algebraic sum* at any moment of the potentials at each lead. The classical limb curves can be derived from the unipolar limb lead curves thus:

Lead I=Lead VL—Lead VR.
Lead III=Lead VF—Lead VL.[2]

(2) Bipolar Chest Leads.—The exploring electrode is placed on any one of the seven chest positions called 1–7; the other electrode is placed on

[1] The lead is from chest position 4 and the right arm; the record closely resembles the unipolar chest Lead VC4 (Fig. 143).

[2] It can be shown that Lead II=Lead I+Lead III.

the right arm (=Leads CR 1 to 7), left arm (=Leads CL 1 to 7) or left leg (=Leads CF 1 to 7).

Origin and Spread of the Cardiac Impulse—The cardiac impulse arises normally in the sino-auricular node, spreads over the wall of the auricles and thus reaches the auriculo-ventricular node. It then passes down the bundle of His, its branches and terminal ramifications, exciting the septum, apex and base of the ventricles in that order. The evidence for these statements is as follows:

Sino-Auricular Node.—(1) The sino-auricular node or *pacemaker*, initiates the cardiac impulse and thus indirectly sets up the heart beat. This node becomes electrically negative before any other part of the auricle, proving that it is the first region to become active.

(2) The P wave represents the invasion of the auricles by the cardiac impulse. Normally, in *bipolar* leads I, II, and III, P is an upward deflection. It can be readily shown that the direction of the deflections recorded by a fixed *pair* of electrodes is related to the direction of spread of the impulse. Thus in the muscle strip illustrated in Fig. 140A, a stimulus applied at S causes the impulse to travel in the direction AB : activity develops first at A, recorded conventionally as an upward deflection, and then at B, recorded as a downward deflection. If the stimulus is applied at S′ the impulse travels in the direction BA ; activity develops first at B, giving an initial downward deflection, and then at A, giving a subsequent upward deflection. Similarly, the normal upward direction of the P wave in the bipolar limb leads is related to the normal pattern of spread of the impulse through the auricles. If the S.A. node is artificially stimulated during diastole, a new impulse is set up there which is associated with a normal P wave because the "forced" (*i.e.* artificially induced) impulse is travelling along the normal route. But if the impulse is set up from any other point on the auricles, the P wave is abnormal in character—flattened or inverted—showing that an unusual path has now been taken by the excitation process (Fig. 166).

When the *unipolar* leads are used, the direction of the P wave depends on whether the proximal or distal zone of the lead becomes negative first. Thus in the unipolar right arm lead (VR) the P wave is normally a *downward* deflection proving that the first region in the auricle to be excited is in the proximal zone of this lead, presumably in the right auricle. In the unipolar left leg lead (VF) the P wave is normally directed upwards, proving that the initial auricular activity is in the distal zone of this lead, presumably in the upper part of the (right) auricle (Fig. 142).

(3) If the S.A. node is extirpated, the P wave in the bipolar limb leads disappears or becomes inverted (Fig. 140B), showing that a new site of impulse formation has been set up. It is worthy of note that the new pacemaker usually proves to be the auriculo-ventricular node (cf. p. 286).

(4) Cooling the S.A. node within physiological limits slows, and warming it quickens, the whole heart. These effects cannot be produced from any other part of the auricle.

Spread in the Auricles.—The impulse spreads out uniformly from the S.A. node over the walls of both auricles as a ripple spreads out over a pond into which a stone has been dropped, or as fluid spreads when poured on to a flat surface. It spreads through the auricular muscle to reach the muscular sleeves surrounding the openings of the great veins, and the auriculo-ventricular

node which lies in the posterior part of the right auricle. No special connecting band exists between the S.A. and A.V. nodes.

Bundle of His.—Experimental evidence proves that the bundle of His is the sole connecting strand between the auricles and ventricles. If the bundle of His is compressed, a proportion of the auricular impulses may fail

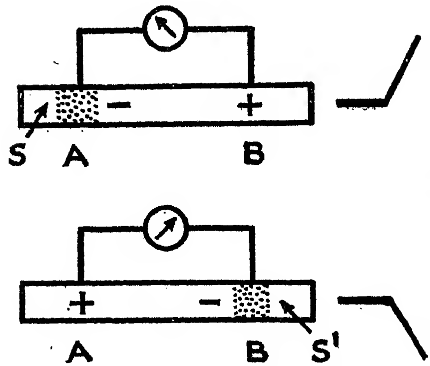


Fig. 140a.—Effect of Direction of Spread of Impulse in Heart Muscle on Direction of Initial Electrical Deflection. Direct Leads.

S : point stimulated on upper strip.
S′: point stimulated on lower strip. The initial deflection is upwards in the upper record, and downwards in the lower record.

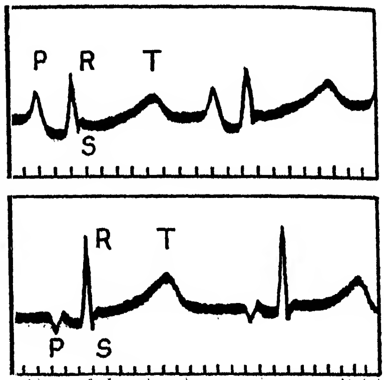


Fig. 140b.—Two Electrocardiograms (cat) taken immediately before (upper record) and after (lower record) crushing the region of the S.A. Node (Lead II). In the lower figure the rate is slower and the auricular wave P is inverted instead of upright, indicating a change of pacemaker. Time in $\frac{1}{30}$ sec. (cf. Figs. 165, 166). (Lewis.)

to reach the ventricles, the rate of which may, for instance, become only half that of the auricles. If the bundle is completely divided in animals or destroyed by disease in man, the auricles and ventricles are completely "dissociated," and beat quite independently one of the other; the auricles then beat at their usual rate in response to impulses generated in the S.A. node, while the ventricles beat much more slowly—thirty to forty times per minute—in response to a new rhythm centre in their own substance, probably situated in the part of the bundle below the point of section (cf. p. 283 and Fig. 162).

Spread in the Ventricles.—The impulse spreads down both branches of the bundle of His to reach the right and left ventricles separately. Fig. 141 illustrates the time relations of the activation of the various parts of the ventricles. The anterior septal region is excited first, whence the impulse spreads to the right anterior and right posterior ventricle. The apex and left posterior ventricle are usually excited later. Lastly the base of the heart is invaded. Note that corresponding points in the right ventricle are affected before those in the left ventricle. It takes the impulse 0·013 second to pass from the bundle of His to the apex, and 0·04 second to arrive at the endocardial surface of the base of the left ventricle. The impulse is carried from apex to base by the Purkinje fibres; injury to the endocardial surface of the

heart beneath which this tissue lies, delays the propagation of the impulse. From each point on the endocardium, the excitation process affects the corresponding part of the epicardial surface. The last part of the heart to become excited is the epicardial surface of the base of the ventricles, after 0·055 second in the case of the right base and after 0·065 second for the left base. It must be noted that the whole of both ventricles is invaded by the excitation wave in well *under* 0·1 *second.*

The figures quoted above are based on results which have been obtained in the dog by the electrical method of timing the arrival of the impulse at various points on the ventricles. Essentially similar results have been obtained by direct observations on the *exposed heart in man.*

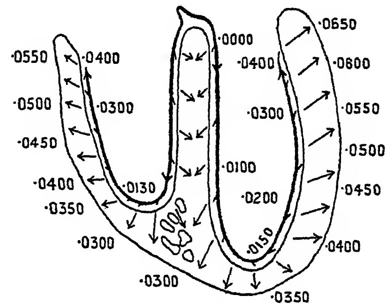


FIG. 141.—Diagram to illustrate the Time Relations of the Spread of the Excitation Process in the Ventricles of the Dog. The figures represent time in seconds; zero time, ·0000, represents the arrival of the excitation process at the top of the interventricular septum. (Lewis.)

If one branch of the bundle is divided (say the right branch), the left ventricle receives its impulse normally, but the impulse reaches the affected right ventricle only after some delay, because it must spread across from the normally stimulated left side to reach the right ventricular muscle and the right branch of the bundle below the side of the block (cf. p. 285).

RATE OF CONDUCTION.—The rate of conduction in the auricular wall is 1 metre per second, in the A.V. node 0·2 metre per second, in the bundle of His and Purkinje fibres 4 metres per second, and in the ventricular wall 0·4 metre per second.

Interpretation of Waves of Human Electrocardiogram.—AURICULAR COMPLEX (P WAVE).—(1) *Limb Leads I, II and III* (Fig. 139).—The first upward deflection P has a rounded or pointed summit, and its duration is 0·1 second. It represents the passage of the impulse from the sino-auricular node over the auricles; the A.V. node is reached at about the *summit* of P. The magnitude of the P wave is some guide to the functional activity of the auricular muscle. In *mitral stenosis* the left auricle is hypertrophied, and correspondingly the P wave is prominent, or bifurcate. In *auricular fibrillation* the P wave disappears and is replaced by a series of fine irregular oscillations, corresponding to the rapid irregular excitation of the auricles (Fig. 172). If the cardiac impulse arises in an *abnormal focus* and spreads in other than the usual direction, the P wave is altered or even inverted, becoming a downward deflection (Fig. 166). The P wave may be prolonged if the spread of the excitation process is delayed as a result of auricular disease.

(2) *Unipolar Leads.*—The features of the P wave in these leads are shown in Figs. 142 and 143. It should be noted that P is normally inverted in Lead VR; in fast records P is seen to consist of several small deflections.

Conduction Time of Bundle of His (P–R Interval).—The time interval from the beginning of P to the commencement of R, *i.e.* the P–R interval, is a guide to the conduction time of the bundle of His. When the classical limb leads are used the *true conduction time* is the interval from the top of P (when the A.V. node is excited) to the beginning of Q (when the invasion of the ventricles commences). It is easier, however, to measure the P–R interval (the Q wave is inconstant and of small size), and undoubtedly it is mainly taken up by the passage of the impulse along the bundle and its branches. The P–R interval in health varies between 0·13 and 0·16 second, the extreme limit of the normal being 0·2 second. When it exceeds 0·2 second there is delayed conductivity in the bundle of His. When the P–R interval is shorter than normal, the impulse has probably arisen in the A.V. node and has therefore excited the ventricles sooner than is normally the case (cf. Fig. 165 and p. 286).

Ventricular Complex (QRST Waves).—Following P there is a brief isoelectric period; a succession of deflections then appears, namely: Q, a small (often inconspicuous) downward one; R, a prominent upward one; and S, another downward one; the record then returns once more to the base line. The duration of QRS in man is about 0·08 sec.; the upper limit of the normal is 0·1–0·12 sec. The final T wave is a broad upwardly directed deflection, with an average duration of 0·27 sec. The duration of QRST is thus about 0·4 sec. The upstroke R just coincides with the onset of ventricular systole; the end of the T wave coincides approximately with the end of ventricular systole or may outlast it slightly.

Interpretation of QRST.—Because of the complex way in which the electrical records are produced when bipolar leads are employed, this question must be considered principally with reference to unipolar leads. Fig. 142 shows the probable distribution in man of the proximal and distal zones of unipolar limb leads VR, VL, and VF, and the corresponding electrocardiograms; the distribution of these zones has been established with considerable accuracy in the dog. As these zones and their related electrocardiograms are similar in dog and man, the results obtained in the dog are probably a useful guide to the events which occur in the human heart. The proximal zone of each lead corresponds to that portion of the heart which is subtended when a solid cone (whose walls are tangential to the heart surface) is projected on it from: (i) the central point in the right shoulder in Lead VR; (ii) the central point in the left shoulder in Lead VL; (iii) a point at the centre of the junction of the torso with the left hind limb in Lead VF.

In all leads QRS represents the stage of invasion, the ST segment and the T wave the stage of repolarization.

(1) Unipolar Right Arm Lead (VR).—In the dog the *proximal* zone of this lead is the anterior right ventricle (*except the right apex*), the right lateral, and the right posterior ventricle, the upper two-thirds of the interventricular septum and the basal paraseptal portion of the anterior left ventricle. The distribution is similar in man. The rest of the heart (except for a narrow intermediate zone) is the distal zone. The initial ventricular deflection in dog and man is a downstroke, Q, which represents negativity starting in the proximal zone. It is followed by an upstroke, R, representing development of relatively greater negativity in the distal zone. There is then a short downstroke, S, representing momentary preponderance in proximal zone

negativity; this is followed by a return to the base line (isoelectric line), representing balanced activity of proximal and distal zone. QRS (in all leads) *corresponds in time and duration with the period of invasion of the*

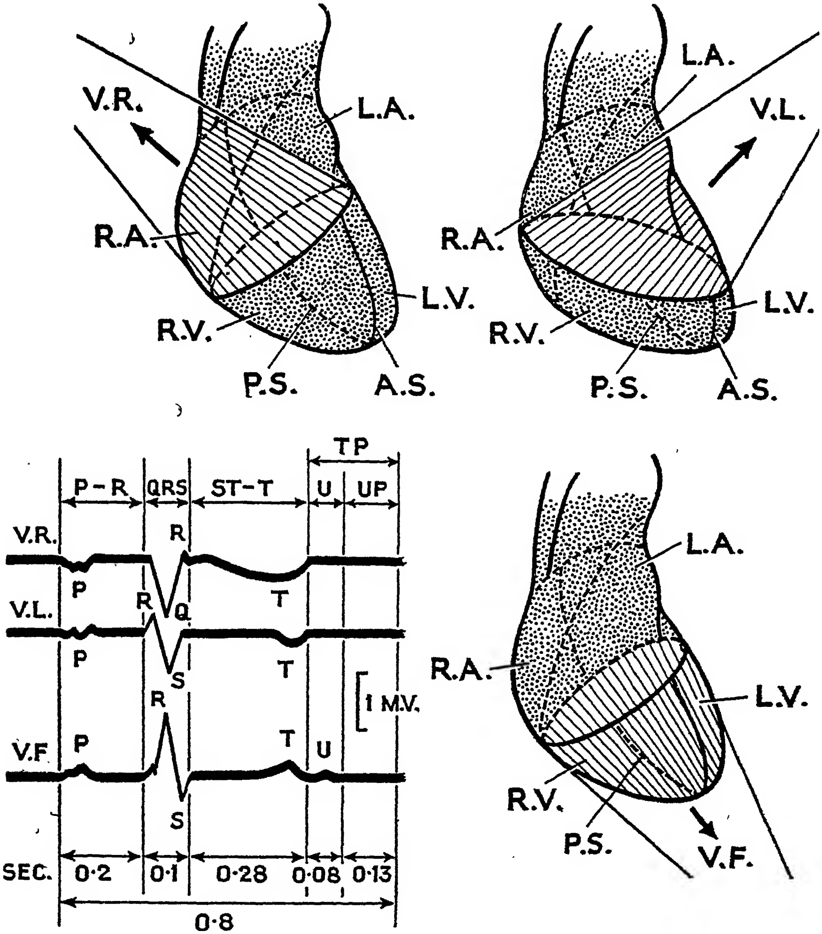


FIG. 142.—Approximate Distribution of Proximal and Distal Zones of Unipolar Limb Leads VR, VL and VF of Human Heart and corresponding Electrocardiograms. (Modified and redrawn from Nahum and Chernoff, in Fulton's *Textbook of Physiology*, W. B. Saunders Co., 1949.)

VR=unipolar right arm lead. VL=unipolar left arm lead. VF=unipolar left leg lead.

ventricles and is the resultant of the complex electrical interference between the proximal and distal zones which occurs during the spread of the invasion process.

The duration of the stage of possession at any point is unknown. The rest of the ventricular complex, *i.e.* the ST segment and T wave, is due predominantly to the complex pattern of repolarization. As relative negativity

in the proximal zone produces a downstroke, relative positivity in this region produces an upstroke; the reverse is the case with the distal zone. In Lead VR, after a short isoelectric period, there is a well-marked inverted (downward-directed) T wave, indicating that repolarization is proceeding more rapidly in the distal than in the proximal zone. Direct experiment shows that the left apical region (which lies in the distal zone of Lead VR) repolarizes first.

(2) Unipolar Left Arm Lead (VL).—In the dog the proximal zone is the upper two-thirds of the anterior surface of both right and left ventricles, and the basal portion of the posterior surface of both ventricles; most of the rest of the heart is in the distal zone. The ventricular complex in Lead VL is readily affected by changes in the position of the heart. When the heart is vertical, VL tends to resemble VR: there is a short initial upstroke R (negativity in distal zone) followed by a big downstroke S (negativity in proximal zone) and a quick upstroke to base line. The T wave is inverted. When the heart is in the normal semi-transverse position, VL tends to resemble VF; there is no Q wave and R is large and upright; the record then returns to the base line; T is upright.

(3) Unipolar Left Leg Lead (VF).—In the dog the proximal zone is the left apex, the major portion of the posterior left ventricle, and the lower third of the anterior left ventricle. There is an initial large upward R wave (initial negativity in the distal zone), then a downstroke S (predominant negativity in the proximal zone) and return to base line. T is erect.

T Wave.—The following general statements may be made about the T wave in any unipolar lead:

(i) A "flat" or absent T wave indicates an equal rate of repolarization of the distal and proximal zones.

(ii) An upright T wave indicates that repolarization is relatively slower in the distal than in the proximal zone.

(iii) A downward T wave indicates that repolarization is relatively slower in the proximal zone.

Deviations from the normal appearance of the T wave in any lead indicate abnormalities in the pattern of the repolarization process. The T wave is readily modified by many conditions, *e.g.* anoxia, ischæmia, application of heat or cold to the heart or by large doses of digitalis, for the reasons just stated.

U Wave.—Nothing is known about the mode of production of the U wave, which is seen in certain leads (*e.g.* VF and some of the chest leads). It occurs early in ventricular *diastole.*

(4) Unipolar Chest Leads.—Fig. 143 shows the proximal zones of the chest positions 1–6 and the corresponding electrocardiograms (VC 1–6). As the exploring electrode is moved across the chest different parts of the heart become respectively the proximal, intermediate, and distal zones with corresponding changes in the configuration of the ventricular complex. On moving from position 1 to 6 the electrical changes in normal man are as follows:

(i) No Q wave is present until position 5 or 6 is reached.

(ii) R increases in amplitude from position 1 to a peak at position 4 or 5.

(iii) The apex of R occurs progressively later from position 1 to 6.

(iv) The S is usually biggest in the right-hand positions and declines or disappears in position 6.

(v) The ST segment (*i.e.* from the end of S to the beginning of T) is approximately isoelectric in positions 1 and 6, but is displaced above the base line in the other positions.

(vi) T is inverted in position 1, is upright in the other positions and is usually maximal in position 4.

(vii) U appears in some leads, usually in positions 2 and 3.

The deflections are to be interpreted in terms of zonal interference of depolarization and repolarization as in the case of the unipolar limb leads.

(5) Classical Limb Leads.—The changes in the QRST deflection that occur in patients have been carefully correlated with the clinico-pathological findings in many disease states and are discussed on pp. 254 *et seq.* and 283 *et seq.* The empirical interpretation is of great value; a theoretical analysis of the curves is difficult for the reasons already given (p. 244).

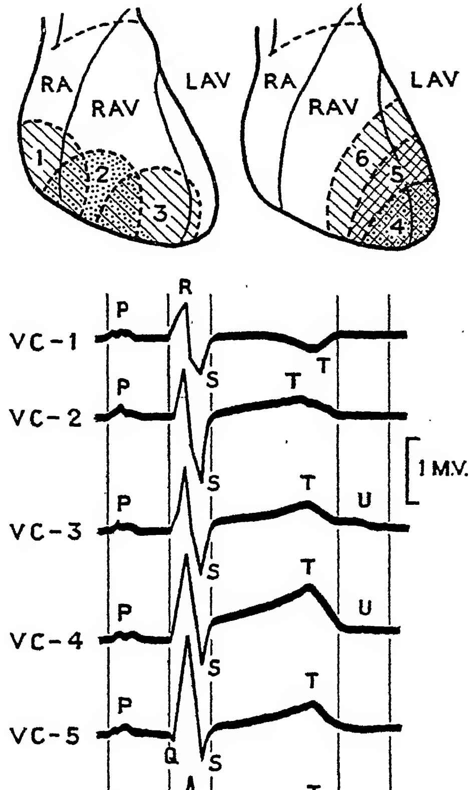


FIG. 143.—Approximate Distribution of Proximal Zones of Unipolar Chest Leads in Human Heart and corresponding Electrocardiograms. (Redrawn from Nahum and Chernoff, in Fulton's *Textbook of Physiology,* W. B. Saunders Co. 1949.)

RA : right auricle.
RAV : right anterior ventricle.
LAV : left anterior ventricle.
On outline of heart areas 1, 2, 3 ; 6, 5, 4 represent proximal zones of corresponding unipolar chest leads VC-1, -2, -3; -6, -5, 4.

Electrocardiographic Changes in Myocardial Lesions. — Lesions of the ventricular myocardium may be produced experimentally (i) by a local application of a 0·2M KCl solution which depolarizes the surface membrane of the affected muscle fibres; (ii) by tying various blood vessels to produce patches of ischæmia and consequent death of the affected areas. Clinically, localized cardiac ischæmia may result from the occlusion of coronary vessels by thrombus or the occurrence of coronary spasm. These myocardial lesions give rise to distinctive electrocardiographic patterns. The initial deflections

of the ventricular complex (representing the stage of invasion) may be comparatively unaltered, but the later phases of the complex are markedly changed. In Fig. 145, (*a*), Lead I, following the peak of R, the downstroke S sets in. Normally, S descends to the base line or below; in this case, however, after a partial descent the record is continued for about 0·2 sec. as an approximately horizontal line (indicated by the arrow) which runs well *above* the normal base line. This abnormality is called *upward deviation of the ST segment.* The record finally descends to the base line and is followed in this instance by a small inverted T wave. In Lead III, the initial deflections R and S are normal; but, after the end of the downstroke S, the record does not rapidly return (as it should) to the base line. There is a partial ascent, after which the record continues for 0·2 sec. as a horizontal line (indicated by the arrow) well *below* the base line; this abnormality is called *downward deviation of the ST segment.* The record finally rises to the base line.

The mode of production of these changes is not fully understood, but the following explanation is usually offered. When a region of the heart is damaged, the affected area generates an *injury potential* (Fig. 144) as is the case with nerve fibres. As a result, the base line of the electrocardiogram during diastole is deflected away from its normal isoelectric position. If unipolar leads are employed, the direction of the deflection is *downwards* when the injured area is in the *proximal* zone of the lead; the direction of the deflection is *upwards* when the injured area is in the *distal* zone of the lead. The magnitude of the deflection is related to the extent of the injured area. During the heart beat following the injury, the stage of invasion ("depolarization") is accompanied by approximately normal electrical deflections. During the stage of possession, however, the injury potential is supposed to be *temporarily suppressed.* After the peak of R or S, the record consequently *returns to, and stays for a short time at, the pre-injury isoelectric level.* The way in which this remarkable change is brought about has not been clearly explained. During the phase of recovery ("repolarization"), the injury potential reappears; the record returns to the initial injury level. The so-called deviation of the ST segment which this record displays is thus attributed to the transient disappearance of the injury potential. The latter part of the ventricular complex may be further modified for the following reason: as the injured area does not undergo repolarization the pattern of repolarization in the ventricles as a whole is altered with resulting changes in the ST segment and the T wave.

Fig. 144 illustrates diagrammatically the changes produced by damage to the ventricular wall. In Fig. 144, Lead I, the injury is presumed to be situated in the proximal zone of the recording leads. The normal base line (N) is deviated *downwards* by the injury potential to the abnormal position I. During the stage of possession, the record temporarily returns to the normal base line N, producing the so-called upward ST deviation. In Fig. 144, Lead III, the injured area is presumed to be situated in the distal zone of the recording leads. The normal base line N is deviated *upwards* by the injury potential. During the stage of possession, the record temporarily descends to the normal level N, producing the so-called downward ST deviation. In Fig. 144, lowest record, similar changes to those shown diagrammatically in Lead I, were produced experimentally by applying a KCl solution to the surface of the cat's ventricle, producing a transient local injury. In such

an experiment, if the post-injury record alone were available, the changes could hardly fail to be misinterpreted. The abnormal base line I would be regarded as the normal isoelectric level.. The transient disappearance of the injury potential would be regarded (as it had been in the earlier clinical studies) as a deviation of the ST segment. In clinical lesions of the heart, the record can hardly ever show the moment of development of the injury. It seems

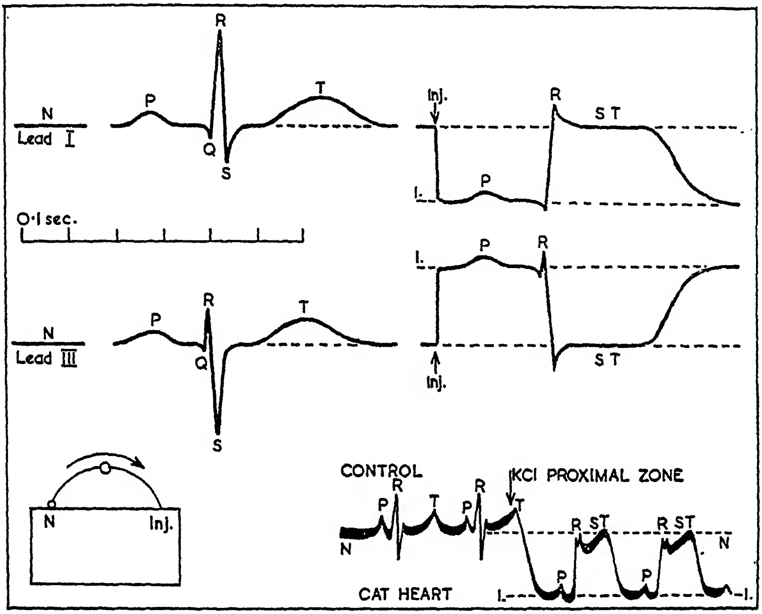


FIG. 144.—Diagram to illustrate the Effects on Electrocardiogram of Localized Injury to Ventricular Wall.

N : Normal base line, *i.e.* true isoelectric level. Time in 0·1 sec.
I : Shift of base line produced by injury potential set up by damaged area.
Upper Record : Lead I. Left hand : Control. Right hand : following lesion in proximal zone of recording leads. Base line deflected downwards.
Middle Record : Lead III. Left hand : Control. Right hand : following lesion in distal zone of recording leads. Base line deflected upwards.
Lower Record : Left hand : diagram to show that injury causes current flow from normal to injured area. Right hand : record showing effect of applying a KCl solution to surface of cat's ventricle. The base line is deflected downwards (lesion in proximal zone of recording lead). This record (modified) is from an experiment by W. F. Floyd and S. Salama.

reasonable to suggest that the changes observed clinically in the ST segment and T wave should be interpreted in the manner indicated above.

The injury potential which produces the so-called ST deviation, and any repolarization changes that may develop, can only be recorded if they occur in the proximal or distal zone of the leads employed. Should they occur in the neutral intermediate zone, they do not affect the recording electrodes. Clinical cardiac lesions generally affect the classical limb leads ; but should the lesion be in the neutral zone for these leads it would not be detectable. In a minority of cases the characteristic changes are only observed in chest

leads or unipolar limb leads. A careful study of many leads also enables the position of the lesion to be localized with considerable accuracy on the basis

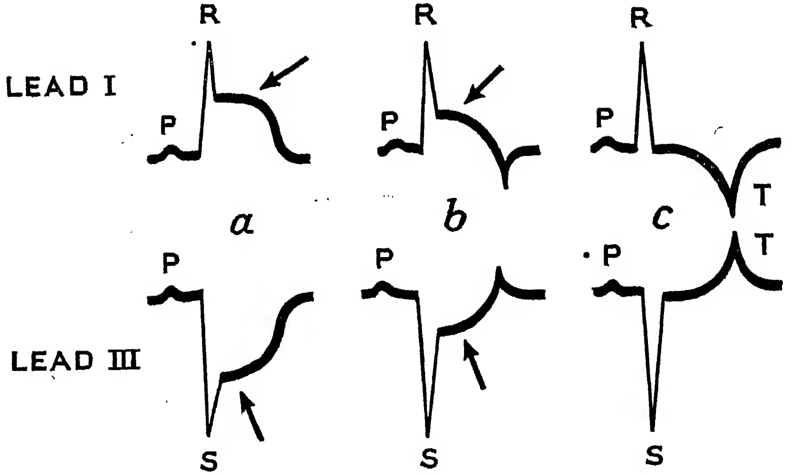


FIG. 145.—Diagram illustrating the Changes in the Ventricular Complexes in Leads I and III. after Coronary Thrombosis affecting Anterior Surface of Ventricles. (Parkinson and Bedford, *Heart*, 1928, *14*.)

(*a*) Elevated ST segment in Lead I and depressed ST segment in Lead III; (*b*) T waves becoming evident; (*c*) ST displacement disappeared. The T wave finally points in the opposite direction to the original ST displacement in each lead.

of the extensive correlations that have been estabished between the distinctive features of the records and the post-mortem findings.

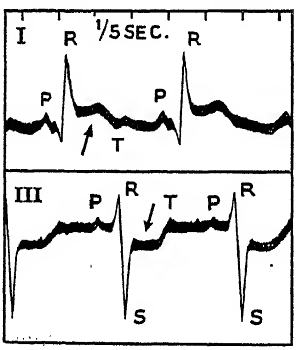


FIG. 146.—Electrocardiogram taken Four Days after the Development of Coronary Thrombosis affecting Anterior Surface of Ventricles.

In Lead I the ST segment (marked by arrow) is elevated above the isoelectric line; in Lead III the ST segment (marked by arrow) is depressed below the isoelectric line.

CLINICAL ISCHÆMIC PATTERNS.—When the left *anterior* descending coronary artery is occluded clinically with resulting ischæmia (infarction) of the anterior surface of the ventricle, curves like those shown in Figs. 142 and 143 are obtained. Soon after the development of the lesion there is upward deviation of the ST segment in Lead I and downward deviation of the ST segment in Lead III; the T wave may disappear. As healing occurs the injury potential is reduced in magnitude and the ST deviation becomes correspondingly smaller. Abnormalities of repolarization persist as shown by abnormalities of the T wave, *e.g.* inversion in Lead I (Fig. 145 (*b*)). After a week the S wave may return to the isoelectric line indicating that the injury potential is no longer produced. The T wave (*i.e.* repolarization) abnormalities persist; in Fig. 145 (*c*) the T wave is inverted in Lead I,

while in both Leads I and III the T wave points in the opposite direction to the original ST deviation. (In some clinical anterior cardiac infarcts there is no ST deviation in Lead I but there is an upward deviation in Lead III.)

If the *posterior* descending coronary becomes occluded clinically, the changes found are the opposite of those shown in Fig. 145; in Lead I the ST deviation is directed downwards and in Lead III upwards.

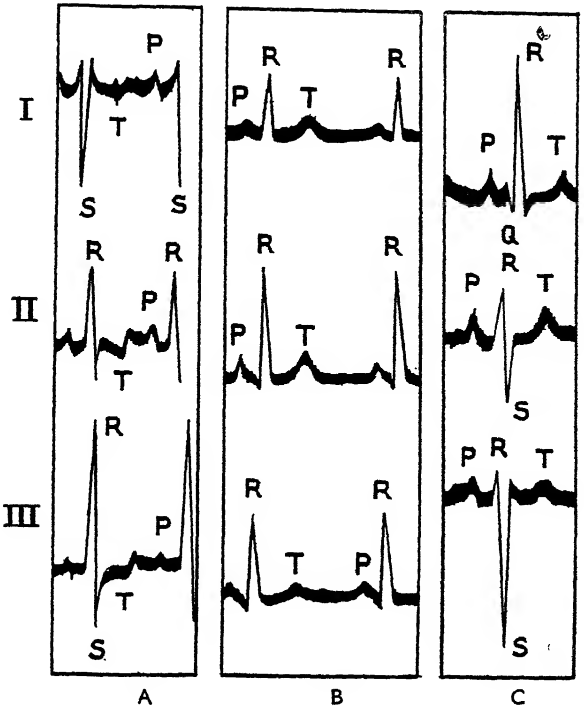


FIG. 147.—Electrocardiogram in Right and Left Axis Deviation.

A. Right axis deviation (in a woman with mitral stenosis). The T wave is inverted in all leads because of treatment with digitalis.
B. Normal record.
C. Left axis deviation (in a woman aged 52 with hypertension; B.P. 245 mm. Hg systolic, 145 mm. diastolic; enlarged left ventricle.

RIGHT AND LEFT AXIS DEVIATION.—Clinically one frequently obtains curves characterized by a tall R wave in Lead I and a big S wave in Lead III. These are called examples of deviation of the main electrical axis of the heart to the left, or briefly *left axis deviation* (Fig. 147, C). Conversely a big S in Lead I and a big R in Lead III is called *right axis deviation.* Such axis deviation is noted with mechanical displacement of the heart or with the movements of the heart which occur with the phases of respiration. Hypertrophy of the *left ventricle,* such as occurs as a result of chronic arterial hypertension or aortic incompetence, is associated with left axis deviation (Fig. 147, C). Conversely hypertrophy of the *right ventricle, e.g.* from

emphysema of the lung or mitral stenosis, is associated with right axis deviation (Fig. 147, A). The whole QRS complex may be prolonged because the greater thickness of the affected ventricle may modify the rate of development of the excitation process.

These changes can be accounted for by reference to the principles set out on p. 248. A displacement of the whole heart or a change in the relation of left to right ventricle alters the position of the proximal and distal zones of certain leads. The deflections in Lead VL when the heart is in the erect position are different from those with the heart in the semi-transverse position (p. 250). It will also be recalled that:

Lead I=Lead VL—Lead VR,
and Lead III=Lead VF—Lead VL.

Characteristic changes would then be expected to occur in Leads I and III; such is actually the case.

The electrocardiographic findings in disorders of the cardiac rhythm are considered on pp. 282 *et seq.*

PRESSURE CHANGES IN THE HEART AND BLOOD VESSELS[1]

Methods Employed.—In the most accurate work, optical methods are employed. In animal experiments a glass tube filled with anticoagulant fluid is introduced directly into the chamber of the heart to be examined. The external end of the tube is closed by a tense rubber membrane upon which a small mirror is fixed. *Pressure* changes in the cavity produce oscillations of the mirror which reflects a beam of light on to a moving photographic plate. The pressure changes in the *right auricle* and the *right ventricle* have been recorded *in man* by replacing the glass tube by a rubber tube which is introduced into an antecubital vein and passed along the venous system to enter the right auricle, or pushed further into the right ventricle (Cournand's method, cf. pp. 260, 279). Indirect information about the right auricular pressure changes in man can be obtained from the *jugular venous pulse* tracing (p. 259).

The *volume* changes in the heart during the phases of the cardiac cycle under varying conditions can be determined by means of a *cardiometer*: this is a glass vessel resembling a thistle funnel which is placed over the ventricles and is fitted round the auriculo-ventricular groove. It is connected with a piston recorder which writes on a moving smoked surface.

Pressure Changes in the Ventricles.—(1) *Isometric Contraction Phase.*—At the onset of ventricular *systole*, the pressures in the auricles and ventricles are approximately the same; the auriculo-ventricular valves are floating into apposition. As the ventricular muscle contracts, the ventricular pressure rises steeply; the A.V. valves[2] are shut, and bulge slightly in a dome-shaped manner into the cavity of the auricles, causing a sudden rise of intra-auricular pressure. No blood can yet leave the ventricle, because the intraventricular pressure is still lower than the intra-aortic; the tension in the ventricle is rising, but no shortening of the muscle occurs yet. This *isometric contraction*

[1] Wiggers, *Pressure Pulses in the Cardiovascular System*, London, 1928.
[2] A.V.=auriculo-ventricular.

phase lasts 0·05 second. Its duration is constant under different circulatory conditions and is not affected by alterations in the heart rate (see Fig. 148).

(2) *Maximum Ejection Phase.*—When the intraventricular pressure rises higher it exceeds the intra-aortic, and the semilunar valves open. Rapid

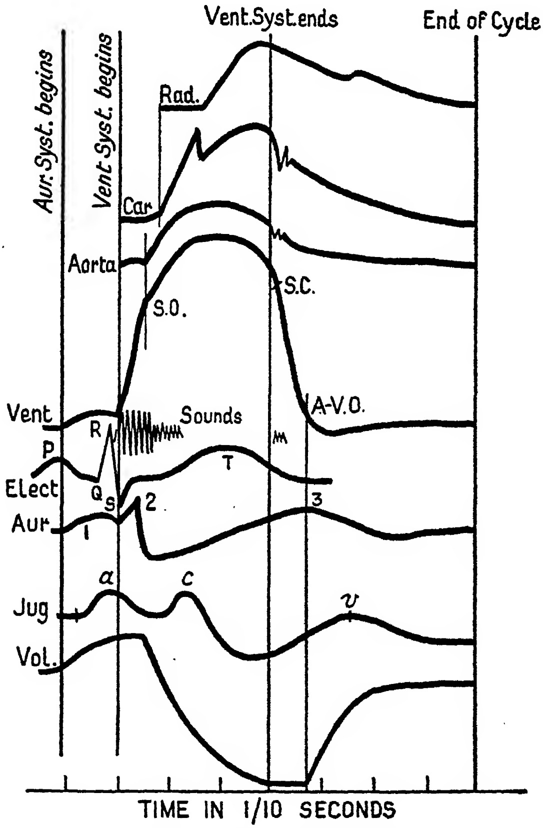


FIG. 148.—Diagram of Pressure, Volume, and Electrical Changes in Heart and Blood Vessels during the Cardiac Cycle. (Modified from Lewis and Wiggers.)

Vol. = Volume curve of left ventricle; Jug. = Jugular venous pulse; Aur. = Intra-auricular pressure; Vent. = Intraventricular pressure; Elect. = Electrocardiogram; Aorta = Intra-aortic pressure tracing; Car. and Rad. = Carotid and radial pulse; S.O., S.C. = Opening and closing of semilunar valves; A.V.O. = Opening of auriculo-ventricular valves. The duration of the cycle is 0·8 second, corresponding to a heart rate of 72 per minute.

ejection of blood occurs, and the volume curve shows that the ventricle is rapidly diminishing in size. As the ventricular muscle shortens, the base of the heart descends and pulls down the auriculo-ventricular ring; the cavity of the auricle is therefore enlarged, causing an abrupt fall of pressure within it. The ventricle and aorta now form one continuous chamber. The aortic pressure thence passively follows the intraventricular, but at a slightly

lower level. In this phase of maximum ejection the pressures in the ventricle and aorta rise to their summit.

(3) *Reduced Ejection Phase.*—The systolic discharge now lessens and little further change in the volume of the ventricle occurs. All parts of the ventricle do not contract for the same length of time; during the latter part of systole some portions of the ventricles cease to contract, and as fewer "units" are functioning the pressure in the ventricles begins to decline slowly. During this period more blood is escaping from the aorta (into the more distal vessels and ultimately into the capillaries) than is reaching it from the heart.

The ejection phase is variable in length, and is chiefly responsible for the variations in the duration of systole. It lasts 0·2–0·25 second. The total duration of systole is best estimated in man by accurately timing the interval between the onset of the first and second sound. When the pulse rate is over 100 per minute, systole lasts less than 0·25 second; between 80–100, about 0·25 second; 65–80, about 0·28–0·3 second; under 65, longer than 0·3 second.

(4) *Ventricular diastole* now sets in, and the intraventricular pressure drops very sharply. A backward flow towards the heart occurs in the aorta, and is halted by the closure of the semilunar valves. The pressure within the ventricle continues to fall steeply as the muscle relaxes, but throughout this period (*isometric relaxation phase*) the ventricle is a closed chamber, and there is no alteration in the length of the muscle fibres.

(5) The auricular pressure which has been rising throughout the greater part of ventricular systole now exceeds the intraventricular. The A.V. valves open because of the difference of pressure between the two chambers, and rapid inflow of blood into the ventricle occurs. It is very important to note that during early diastole a large percentage of the total ventricular filling —*i.e.* about 60%—takes place in this way; a good deal of blood can thus enter the ventricles in the *absence of any effective auricular contraction* as in *auricular fibrillation* (p. 293). As the auricular and ventricular pressures become equal, little further inflow into the ventricle occurs. When the diastolic pause is a long one, very little additional increase in ventricular volume takes place (phase of *diastasis*).

(6) *Auricular systole* now sets in, and the contents of the auricle are driven into the ventricle. On an average, auricular systole only contributes 35% of the total ventricular output. The exact proportion is dependent on:

(i) Time in diastole at which auricular systole occurs.
(ii) Vigour of auricular systole.
(iii) Completeness with which the ventricle is already filled.

If the auricle contracts after a *short* diastole it may contribute a considerable amount of blood, up to 60% of the ventricular output. If it contracts at the end of a *long* diastole, the ventricle may already be so full that the auricle can contribute relatively little. Auricular systole lasts about 0·1 second.

Pressure Changes in the Auricles.—The pressure changes in the auricles may be briefly summarized (Fig. 148, Aur.):

(1) The *first* positive wave is due to auricular systole. All the fibres of the auricle do *not* contract simultaneously—"fractionate contraction" occurs; the fibres nearest to the S.A. node contract first, and so the intra-auricular pressure begins to rise. At the height of the wave most of the auricular fibres are in action. Late in auricular systole relaxation has begun in

most of the fibres; only a few are still contracting, and so the pressure falls.

(2) The *second* positive wave is due to the bulging of the A.V. valves which occurs at the onset of ventricular systole (p. 256).

(3) A sudden fall of pressure then occurs. This is due to:

(i) The negative intrapleural pressure, which pulls on the relaxed auricle and enlarges its cavity.

(ii) The ventricular muscle shortens, pulls down the A.V. ring (p. 257), and also helps to enlarge the cavity of the auricle, so that the pressure within it falls.

(4) The pressure rises (*third* positive wave) as blood accumulates in the auricle while it is still shut off from the ventricles by the closed A.V. valves. This wave continues beyond the end of ventricular systole to the end of the isometric relaxation phase. When the ventricular muscle relaxes, the A.V. ring moves up again. The cavity of the auricle is thus made smaller, and this helps further to raise the pressure within it. The A.V. valves now open, the auricle empties itself into the ventricle, and the pressure falls rapidly.

The behaviour of the auricle during the cardiac cycle may be compared with that of a concertina. During ventricular systole its cavity is enlarged by the negative pressure in the thorax and the downward movement of the A.V. ring; with the onset of diastole it is compressed by the ascent of the A.V. ring.

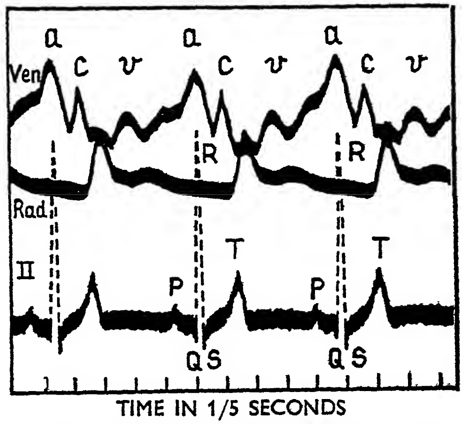


FIG. 149.—Simultaneous Normal Venous, Radial, and Electrocardiographic Curves. (Lewis.)

Delay in transmission through the air of the rubber tubing displaces the venous and radial records to the right by approximately 0·03 of a second.

Jugular Pressure Tracing.—A good idea of the pressure changes in the *right auricle in man* can be obtained by recording the pressure changes in the jugular vein. The right vein is employed, because it is in a direct line with the superior vena cava and right auricle. A small metal cup is placed over the vein and connected with a sensitive Marey's tambour. By means of Mackenzie's polygraph, simultaneous jugular and radial tracings can be obtained (Fig. 149).

The jugular tracing shows three positive waves, called *a*, *c*, *v*, respectively.

a Wave.—The *a* wave is due to auricular systole. As the pressure in the auricle rises it becomes more difficult for the jugular vein to empty itself, so that the pressure within it is increased. As already mentioned, there is a sleeve of auricular muscle which surrounds the opening of the great veins, and closes the orifice of the superior vena cava during auricular systole by a sphincter-like action. *No* regurgitation of blood into the great veins occurs

during auricular systole. The pressure within them rises because of stasis of the blood stream, producing the *a* wave.

c Wave.—The *c* wave is due to the onset of ventricular systole causing bulging of the A.V. valves into the auricle and raising the pressure there. The *a–c* interval, *i.e.* the interval between the beginning of the *a* and *c* waves respectively, is a guide to the conduction time of the bundle of His. More exactly it is the interval between the onset of auricular and ventricular systole. If the *a–c* interval exceeds 0·2 second, delayed conductivity of the bundle of His is present (Fig. 161, p. 283).

v Wave—Like the third positive wave in the auricle, the *v* wave is due to (i) the filling of the auricle while the A.V. valves are shut, and (ii) the upward movement of the A.V. ring which occurs at the end of ventricular systole. The summit of the *v* wave marks approximately the end of ventricular systole.

The negative waves are labelled x, x', and y, and are caused in the same way as the depressions in the auricular pressure tracing.

The events in the jugular vein naturally occur slightly later than the corresponding events in the auricle. The onset of the *c* wave precedes the primary wave in the radial artery by 0·1 second; the top of the *v* wave precedes the dicrotic notch by a similar interval (see Fig. 149).

The pressure changes in the jugular vein are of importance in giving information about the human heart that could otherwise only be obtained electrocardiographically. The *a* wave is an indication of auricular systole. It is absent in auricular fibrillation and flutter, when the normal co-ordinated auricular contraction is lost. An auricular premature contraction is associated with an early *a* wave (Fig. 166); a ventricular extrasystole may coincide in time with the normal auricular contraction and is therefore not preceded by an *a* wave (Fig. 167). The *a–c* interval is an index of conductivity in the bundle of His. It is prolonged in bundle lesions (*supra*); it is less than normal when the impulse arises in the A.V. node.

Human Right Ventricular Pressure Curves.[1]—These curves resemble in their general outlines those obtained experimentally in animals (Fig. 150). The pressure during diastole is about atmospheric, *e.g.* 2 mm. Hg, and it rises to 20–25 mm. Hg at the height of systole; there is some fluctuation in the pressure with the phases of respiration. When there is obstruction in the pulmonary circuit the resistance to the flow of blood through the lungs is increased and as a result the right ventricle must contract more vigorously; right ventricular pressure consequently rises Such changes occur in pulmonary fibrosis and in the later stages of congestive heart failure (cf. p. 460). When the right ventricle begins to fail it does not empty itself completely during systole and consequently ventricular *diastolic* pressure rises; in one case studied (Fig. 150), right ventricular pressure was about 20 mm. Hg during diastole and 80–100 mm. Hg at the height of systole; when clinical improvement set in the curves became more normal.

Aortic Pressure Changes.—(1) The semilunar valves open 0·05 second after the onset of ventricular systole. Blood now passes from the ventricle into the aorta, and the aortic pressure rises smoothly to a maximum. During the latter part of systole the pressure in the aorta falls slightly: the ventricle is now in its reduced ejection phase, is contracting less forcibly and is expelling relatively little blood. The inflow into the aorta from the heart is now less

[1] Cournand *et al.*, *Proc. Soc. exp. Biol. Med.*, 1944, *55*, 34.

than the escape from the aorta and arteries into the capillary bed. It is possible that through the sino-aortic nerves the peripheral resistance is reflexly diminished towards the end of systole, aiding the outflow of blood from the arterial system (Fig. 148).

(2) When ventricular diastole sets in, the pressure in the ventricle drops sharply. This causes a backward flow in the aorta together with a similar drop in aortic pressure termed the *incisura*. The fall of aortic pressure is halted by the closure of the semilunar valves, which causes a slight rebound of the column of blood and is responsible for small after-vibrations. Another factor to be considered is that the aorta, which has been stretched during systole, now recoils. Some of the blood contained within it is driven back

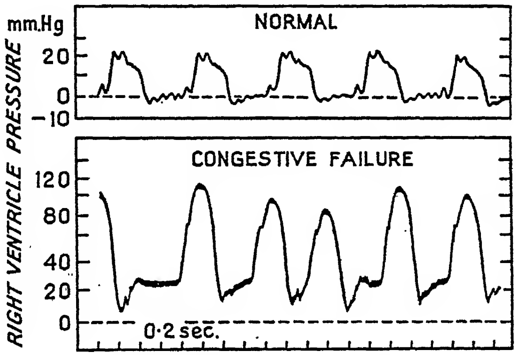


FIG. 150.—Right Ventricular Pressure Curves in Man, (Cournand *et al. Proc. Soc. exp. Biol. Med.*, 1944. *55*, 34.)

Upper Record: Normal. One respiratory cycle is shown.
Lower Record: Case of congestive heart failure; mitral stenosis, aortic incompetence, auricular fibrillation. Note raised diastolic pressure (associated with increased right auricular and peripheral venous pressure). Marked elevation of right ventricular systolic pressure. Cardiac output subnormal; total blood volume twice normal; systemic B.P. 210/90. Note respiratory variations associated with dyspnœa. Time in 0·2 sec.

towards the heart and helps to close the valves, while some is driven on towards the periphery and helps to set up a second positive wave in the arterial system at the beginning of diastole.

The aortic pressure slowly falls throughout the rest of diastole as the leak through the arterioles into the capillaries continues.

Arterial Pulse.—The blood which is suddenly thrown into the aorta during systole is accommodated partly by moving the entire arterial column on at greater velocity, and partly by stretching the arterial wall. This *increase of pressure* and this *arterial distension* are transmitted from one segment of an artery to the next, in the form of a wave—the *pulse wave*—which is independent of the velocity of the blood flow. *The rate of the blood flow varies inversely as the total cross-section of the vascular bed in any part:* it is 0·8–1·0 metre per second in the aorta and in its larger branches, considerably less in the arterioles and only about 0·5–1·0 *millimetre* per second in the capillaries. It increases steadily in the veins the nearer one gets to the heart, owing to the decrease in their total cross-section. The *velocity of the*

pulse wave is considerably *more rapid* than that of the blood. The average pulse wave velocity in metres per second at different ages is: age 5, 5·2; age 20, 6·2; age 40, 7·2; age 70, 8·3, with a range of 1 metre above or below these figures; *i.e.* the velocity increases as age increases.[1]

These results enable us to determine the effect of age on the elasticity of arteries. By *elasticity* of an artery is meant the *percentage increase in the volume of the artery with each mm. Hg increase in blood pressure.* It is related to the velocity of the pulse wave thus:

$$\text{Elasticity} = \frac{1{\cdot}72}{(\text{velocity of pulse wave})}$$

From these figures the elasticity at different ages can be calculated: age 5, 0·47; age 20, 0·33; age 40, 0·24; age 70, 0·18; there is thus a *notable decrease in elasticity as age advances.* This means that with the same cardiac output the *blood pressure is raised during systole more in older people than in younger* as less fluid can be accommodated by distension of the arterial system.

Radial Pulse.—Two waves are present—the *primary* wave and the *dicrotic* wave (Fig. 149). In the radial artery the upstroke occurs 0·1 second later than that in the carotid. The tops of the waves are rounded and their amplitude is smaller. The end of systole is approximately indicated by a rounded *dicrotic* dip or notch. The *primary* wave (or percussion wave) in the radial first rises and then falls towards the end of systole (as does the corresponding wave in the large arteries). The initial rise (expansion of the artery) is due to the fact that blood is pumped by the heart into the blood vessels more rapidly than it escapes into the capillaries; the later fall is due to the reduced ejection phase of the ventricle delivering less blood than is passing out through the arterioles (which may be reflexly relaxed). When the dicrotic notch is thus present on the downstroke, the pulse is called *catacrotic* in character.[2] The *dicrotic wave* which follows the notch represents a modification of the vibrations set up by the closure of the aortic valves and the rebound of the blood from these curtains. (See also Figs. 161 *et seq.*)

The *details* of the pulse form can only be studied in optically recorded tracings (*e.g.* as in Fig. 149). For most clinical purposes records taken with *Dudgeon's sphygmograph* suffice.

A large primary wave may be due to (i) large output; (ii) slow heart rate; (iii) low peripheral resistance. A small primary wave is due to (i) small output; (ii) high peripheral resistance from increased arterial tone or (iii) impaired elasticity of the vessels.

It can be shown in a circulatory model that the descending limb of the pulse curve falls more steeply if the *diastolic pressure is low*, the *arterial wall thick*, or if *aortic regurgitation is present.* Similarly, after injection of nitrites, stimulation of the aortic nerve to produce relaxation of the arterioles, and in aortic disease in man (p. 298), a rapid downstroke is observed.

Circulation Time.—By this is meant the time taken for the blood to travel from one point in the circulation to another. It is thus a measure of the average *linear velocity* of the blood flow which is *directly related to the*

[1] Bramwell, Hill, and MacSwiney, *Heart*, 1923, *10*, 233.

[2] Occasionally a break or notch occurs on the ascent, which is termed the *anacrotic* notch. Its significance is not yet clearly understood. It is observed when the aorta is compressed, in aortic stenosis and in arteriosclerosis.

cardiac output. Thus a rise or fall in the cardiac output tends to produce a decrease and increase respectively in the circulation time. The time can also be affected by local changes in the vascular bed, *e.g.* in the systemic veins or lungs. The principle of the methods employed is to inject some substance into a systemic vein (unless otherwise stated an ante-cubital vein is used) and to determine the time of its arrival at some point in the systemic arterial system or in the lungs.

The substances commonly used are:

(1) *Decholin* (Na salt of dehydrocholic acid).—5 c.c. of a 2% solution are injected rapidly into an ante-cubital vein. When the drug reaches the mouth and pharynx it gives rise to an intensely bitter taste and smell. The "arm to tongue" circulation time is the interval which elapses between the beginning of the injection to the first recognition of the bitter taste. The taste persists for 15–20 seconds.

(2) *Sodium Cyanide.*[1]—0·1 mg. of NaCN per kg. body-weight is injected intravenously. On reaching the carotid body (and the aortic body) the chemoreceptors are stimulated and send up excitatory impulses to the respiratory centre. The end point consists of dilatation of the alæ nasi followed by a deep inspiratory gasp and 6–10 rapid breaths. This gives the "arm to carotid bifurcation" time.

(3) *Magnesium Sulphate.*[2]—6 c.c. of a 10% solution of $MgSO_4$ are injected intravenously. On reaching the throat there is a sudden start combined with an intense feeling of heat in the pharynx; the heat is progressively felt in the face, hands, and finally in the feet. The sensation passes off in 10–20 seconds. This method gives the "arm to hand" and "arm to foot" time.

(4) *Ether.*[3]—5 minims of ether (in an equal volume of saline) are injected intravenously. When the ether reaches the respiratory passages it produces a gasp, cough, and a grimace, and ether can be smelt in the breath. This method gives the "arm to lung" time.

The normal range of circulation time (under basal conditions) by the above methods in man is:

Decholin, 14–19 seconds.	Magnesium Sulphate, 6–16 seconds.
Sodium Cyanide, 12–33 seconds.	Ether, 3–9 seconds.

Clinical Value of Circulation Time Determination.—(1) *Metabolic Rate.*—There is a significant correlation between the basal metabolic rate (B.M.R.) and the circulation time; subjects with the highest B.M.R. have the shortest circulation time. Thus cases of hyperthyroidism with a raised B.M.R. have a circulation time (decholin) which is below normal limits, *e.g.* 9·4–11·8 seconds. When the B.M.R. is lowered by administering iodine or by partial thyroidectomy the circulation time rises. Thus one case of Graves' disease with a B.M.R. of +20 had a circulation time of 12·2 seconds: following thyroidectomy the B.M.R. fell to —28 and the circulation time rose to 20 seconds.

(2) *Decompensated Heart Disease.*—The circulation time is usually unchanged in compensated heart disease. It tends to be prolonged as decom-

[1] Stead and Kunkel, *Amer. J. med. Sci.*, 1939, *98*, 49. Smith and Allen, *Proc. Staff Mayo Clin.*, 1941, *16*, 53.
[2] Bernstein and Simpkins, *Amer. Heart J.*, 1939, *17*, 218.
[3] Baer and Isard, *Amer. J. med. Sci.*, 1940, *200*, 209.

pensation develops (though it may be normal even in advanced cases of decompensation). A single determination is of limited value, but repeated observations in the same individual may be a useful guide to the course of the case. Fig. 151 shows the range of circulation time and the average values in different grades of decompensation. A normal ether time argues against the presence of right heart failure; if the ether time is normal and (say) the decholin time is prolonged, there is probably slowed circulation in the lungs (indicative usually of pulmonary congestion).

(3) *State of Peripheral Vessels.*—Variations in the circulation in the systemic veins due to local conditions may affect the circulation time.

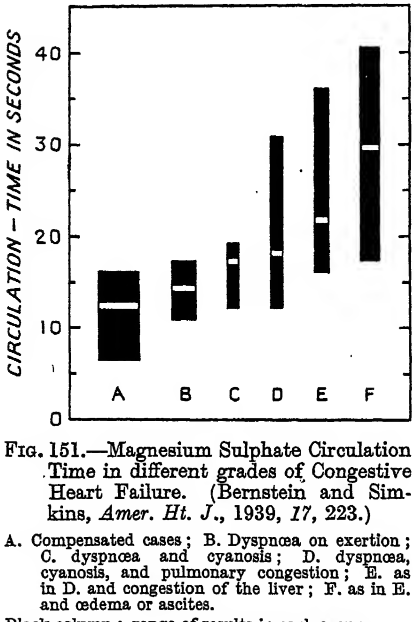


FIG. 151.—Magnesium Sulphate Circulation Time in different grades of Congestive Heart Failure. (Bernstein and Simkins, *Amer. Ht. J.*, 1939, *17*, 223.)

A. Compensated cases; B. Dyspnœa on exertion; C. dyspnœa and cyanosis; D. dyspnœa, cyanosis, and pulmonary congestion; E. as in D. and congestion of the liver; F. as in E. and œdema or ascites.

Black column: range of results in each group.
White bar: mean of observations in each group.

(i) With the hand (of the arm used for the intravenous injection) placed in water at 27° C., the NaCN "arm to carotid" time was 47 seconds; when placed in water at 40° C. the time was 27 seconds. The difference is due to slowed circulation in the arm veins produced by local cold in the former case and the quickened circulation produced by local heat in the latter case and is no guide to the state of the cardiac output.

(ii) The "foot to carotid" time is much longer than the "arm to carotid" time (38·7 seconds against 20·1 seconds in observations on about 80 subjects). This difference is a measure of the extra time taken by the venous blood from the foot to return to the heart. The "foot to carotid" time is shortened by: (*a*) elevation of the legs (gravity promoting the venous return); (*b*) sympathectomy (arteriolar dilatation speeding up the limb blood flow); (*c*) a short bout of leg movements in the supine position (because of the resulting local vasodilatations and compression of the veins); (*d*) alterations in skin temperature. Following abdominal *operations* the "foot to carotid" time is progressively prolonged and may reach a maximum of 50% above the pre-operative level after 5–10 days. (The "arm to carotid" time is little altered.) This considerable slowing down of the circulation in the legs occurs about the time when the incidence of post-operative *thrombosis* in the leg veins is maximal (cf. p. 149).

Movements of the Heart.—The heart is enclosed within the pericardium, which is fixed above to the great vessels and below to the central tendon of the diaphragm. The fibres of the heart are arranged in a complex spiral manner.

(1) During systole, all the diameters of the heart are decreased and the position of the organ alters. The base and auriculo-ventricular ring *descend*,

the apex of the heart is rotated anteriorly and to the right. It is thus brought into closer apposition with the chest wall, and this impact mainly accounts for the *apex beat*, which is the shock felt in the fifth left intercostal space $3\frac{1}{2}$ inches from the middle line.

(2) The entire heart moves downwards during inspiration, and the apex is rotated clockwise. The shadow of the heart, as seen during X-ray examination, lengthens and becomes narrower.

(3) The position of the heart alters with that of the body. The heart is displaced to the left and pressed more firmly against the chest wall when the subject lies in the left lateral position. Anything which pushes the diaphragm up—*e.g.* assuming the sitting or the prone position, or flatulence, causes the heart to lie more horizontally.

Heart Sounds.[1]—The sounds are usually heard clinically with the aid of a stethoscope; for purposes of refined study they may be recorded by applying a microphone to the chest and connecting it through a suitably arranged circuit to an oscillograph. The interpretation of these records is facilitated by a simultaneously taken jugular pressure tracing and electrocardiogram. The classical first and second heart sounds can be heard easily and regularly. The so-called third and auricular sounds are only heard with great difficulty; in most normal subjects they cannot be detected by routine auscultatory methods. They are, however, regular features of sensitive graphic records (Fig. 152).

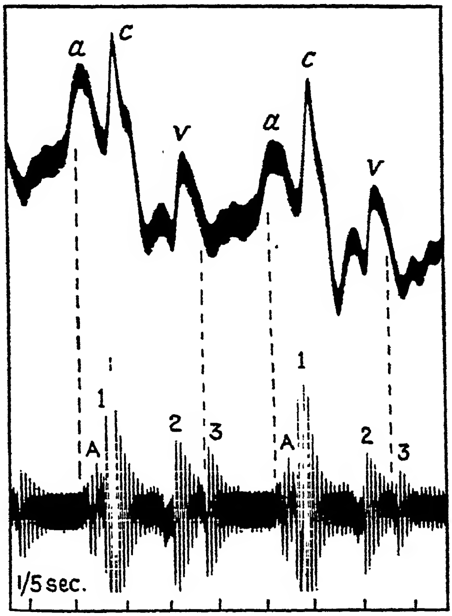


FIG. 152.—The Four Normal Heart Sounds. (Orias and Braun-Menendez, *The Heart Sounds*, London, 1939.)

Records from above downwards: jugular venous pressure tracing; heart sounds; time in $\frac{1}{5}$ second. In the heart sound record, 1, 2, 3 represent the first, second, and third sounds and A, the auricular sound.

FIRST SOUND.—The first sound is prolonged and low pitched. It coincides with the spike of the R wave of the electrocardiogram and just precedes the onset of the *c* wave of the venous tracing. Its duration is 0·1–0·17 second (Fig. 153). It is undoubtedly due to *ventricular systole*; it sets in with the commencement of ventricular contraction but ends before systole is over. Several mechanical factors are responsible for the first sound: (i) contraction of the ventricular muscle; (ii) sudden closure and tension of the auriculo-ventricular valves; (iii) vibration of the structures in the mediastinum and chest wall produced by the movement of the heart and its impact against

[1] Orias and Braun-Menendez, *The Heart Sounds*, London, 1939. Rappaport and Sprague, *Amer. Heart J.*, 1941, *21*, 257.

the chest. Frequently the first sound consists of two distinct groups of vibrations corresponding to (*a*) the isometric contraction phase, and (*b*) the ejection phase. The presence of the latter group suggests that vibrations set up by the movement of the blood are also a contributory factor. Most of the vibrations have a frequency of 25–45 per second. The *intensity* of the first sound is not related to the cardiac output but to the tension developed during the isometric contraction phase.

Second Sound.—This is of higher pitch, abrupt and clear, and is best heard at the base of the heart. It occurs precisely at the onset of *diastole* and is due to closure of the semilunar valves in the aorta and pulmonary artery. Its intensity varies with the blood pressure in the great vessels at the *onset* of diastole. The duration of the second sound is 0·1–0·14 second ;

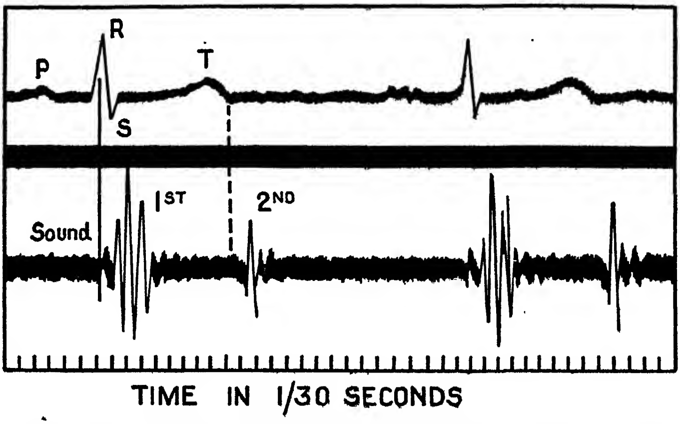


Fig. 153.—Simultaneous Electrocardiogram and Heart Sound Curve from a Normal Man.

The figure shows the time relations of the electrocardiogram to the beginnings of the 1st and 2nd heart sounds (other heart sounds are not shown). Points on the vertical line are simultaneous. (Lewis.)

it is composed of 3 or 4 principal vibrations with a frequency of 50 per second. It corresponds to the notch which is constantly present on the upstroke of the *v* wave in jugular tracings in optical records (Fig. 152) ; it may precede (Fig. 148), coincide with (Fig. 153), or follow the end of the T wave.

Third Sound.—Some fortunate observers claim to hear this sound by auscultation in a considerable proportion of normal subjects (*e.g.* 60%). Its duration is about 0·1 second ; it coincides with the last portion of the descending limb of the *v* wave, *i.e.* with the end of the phase of rapid filling of the ventricles which takes place when the auriculo-ventricular valves open early in diastole. It is due to this inrush of the blood into the ventricles setting up a series of vibrations (Fig. 152).

Auricular (Fourth) Sound.—This is graphically recorded as a complex group of vibrations commencing with the rise of the *a* wave of the venous pulse. It is rarely heard on auscultation in normal people.

The auricular sound is well recorded graphically in cases of complete heart block (Fig. 154). The sound then follows regularly on each P wave (which represents auricular excitation) and shows up clearly when

separated by a long interval from the sounds resulting from ventricular activity.[1]

CARDIAC MURMURS.—These are generally produced when the blood flows through a narrowed orifice from one wider channel into another. The murmur may be:

(i) *Systolic: e.g.* when blood (*a*) leaks back during ventricular systole from the ventricle into the auricle because of incompetence of the auriculo-ventricular valves (Fig. 173); (*b*) flows during systole from the ventricle through a narrowed aortic or pulmonary opening into the corresponding artery; (*c*) flows during systole from the left into the right ventricle through a patent inter-ventricular septum.

(ii) *Diastolic: e.g.* when the blood (*a*) leaks back from the aorta into the left ventricle during diastole because of incompetence of the aortic valves (Fig. 174), or (*b*) leaks passively in cases of mitral stenosis from the overfilled left auricle into the left ventricle during diastole of the heart because of the high intra-auricular pressure (p. 258).

(iii) *Presystolic* (*i.e.* coinciding with auricular systole): this murmur occurs characteristically when the left auricle during systole is driving its contents through a narrowed mitral orifice (mitral stenosis) into the left ventricle (p. 297).

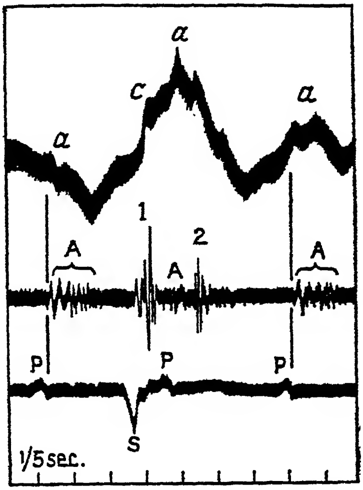


FIG. 154.—Heart Sounds in Complete Heart Block. (Orias and Braun-Menendez, *The Heart Sounds*, London, 1939.)

Records from above downwards are: jugular venous pulse, heart sounds, and electrocardiogram. Note that each P wave is followed by an auricular sound (A) and by an "*a*" wave in the jugular pulse. The ventricular complex (mainly an S wave) is followed by the first and second heart sounds (1, 2).

(iv) *Systolic and Diastolic: e.g.* when the blood flows throughout the cardiac cycle from the aorta into the pulmonary artery through a patent ductus arteriosus (p. 333).

[1] *Reduplication of Heart Sounds.*—By reduplication of a sound is usually meant splitting of its component vibrations into two groups.

(1) *Of first sound.*—(*a*) This may occur physiologically as an exaggeration of the separation normally existent between the two constituent groups of vibrations; (*b*) in bundle branch block when there is partial dissociation and asynchrony of the contraction of the two ventricles (very rarely).

True reduplication may be due to an *audible auricular sound* as when the auricle is hypertrophied or the P–R interval is unduly prolonged.

(2) *Of second sound.*—Splitting may be due to the asynchronous closure of the aortic and pulmonary valves (*e.g.* in branch bundle block).

True reduplication may be due to a *louder third sound*; if mitral stenosis is present it may be caused by a sound produced when the diseased valves are opened by the flow of blood into the ventricle in early diastole ("opening snap" of mitral valves).

The term "*triple heart rhythm*" (*gallop* rhythm) describes the cadence produced when *three loud sounds* recur in successive cardiac cycles. It may occur in healthy young subjects with a loud third sound. For its significance in disease see Evans, *Brit. Heart J.*, 1943, *5*, 205.

REGULATION OF THE HEART RATE[1]

NERVE SUPPLY OF THE HEART (cf. pp. 709, 712).

1. **The Vagus.**—The connector fibres arise in the dorsal nucleus of the vagus and run in the vagus trunk to end in nerve cells in the sino-auricular and auriculo-ventricular nodes. The right vagus chiefly supplies the S.A. node, and the left vagus ends mainly in the A.V. node. From the nerve cells present in the two nodes, a second relay of fibres arises to supply the auricles, the bundle of His, and the base of the ventricles. The apex of the ventricles receives no vagal fibres.

The vagus exerts a continuous restraining action on the rate of the heart (*tonic inhibitory action*). This is proved by the acceleration which follows section of these nerves in most animals, or " paralysis " of the nerve endings by atropine in man. The degree of vagal activity, or *vagal tone* (*vagal restraint*) as it is called, is the chief factor controlling the rate of the human heart at rest.

Stimulation of the peripheral end of the cut vagus produces the following effects which prove that the nerve can depress every part of the cardiac mechanism (Fig. 155).

(i) *S.A. Node.*—The rate at which impulses are elaborated here is diminished, and so the whole heart is slowed (right vagus).

(ii) *Auricle.*—There is considerable diminution in the force of contraction of both auricles. The duration of systole is greatly shortened, and so the refractory period is correspondingly diminished.

(iii) *Bundle of His.*—Conductivity is impaired to a varying extent. First the P–R interval of the electrocardiogram is prolonged; then occasional impulses from the auricles fail to be transmitted through the bundle, and a 2 : 1 or other regular relationship between the auricular and ventricular rates is established. Conductivity may be entirely abolished, and auricles and ventricles thus become completely dissociated. As already noted, the ventricles are so " accustomed " to being driven from the auricles that, should their functional connections with those chambers be suddenly destroyed, they remain quiescent for a brief period and do not beat at all. In this way the vagus may produce temporary ventricular silence. The ventricles as a rule soon resume beating at their own slower independent rate; this " idioventricular rhythm " is outside nervous control altogether. Block of one or other branch of the bundle of His may be produced, with characteristic alterations of the electrocardiographic record (p. 285).

(iv) *Ventricle.*—There is diminution in the force of contraction of the ventricles; the rate is usually markedly decreased, and the rhythm may be disturbed.

The vagus thus—(α) slows the whole heart; (β) produces varying degrees of heart block, or arrests the ventricles altogether; (γ) diminishes the force of contraction of the chambers of the heart; (δ) the duration of systole (and therefore of the refractory period) is shortened.

(v) *Effect on the Blood Pressure.*—As the heart is slowed and its force lessened, the *output of the heart is greatly reduced*, and the blood pressure falls rapidly (Fig. 155). Within a few seconds of commencing vagal stimula-

[1] Boas and Goldschmidt, *The Heart Rate*, Springfield, Ill., 1932.

tion the blood pressure may sink almost to zero, and it remains low. The blood accumulates on the venous side of the circulation—the right auricle and great veins—and the venous pressure rises. The arterial flow into the organs is greatly reduced because of the fall in cardiac output, but (for a short period of stimulation) the venous outflow from the organs is less

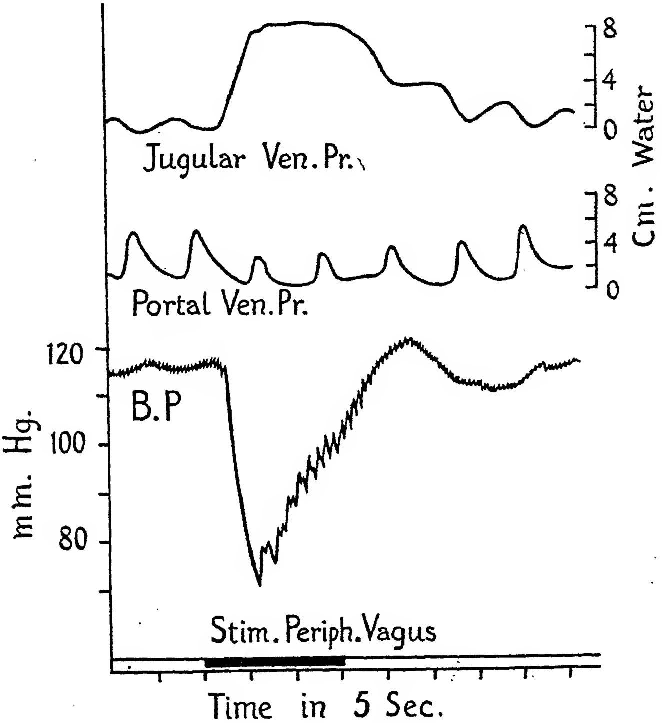


FIG. 155.—Effects of Peripheral Vagal Stimulation on Arterial Blood Pressure, Jugular and Portal Venous Pressures.

Experiment on Cat. Arterial B.P. in mm. Hg; jugular and portal venous pressure in cm. H_2O. (Excursions on venous tracings are due to respiration.)
Note great fall of arterial blood pressure and marked slowing of heart; rise of jugular venous pressure; fall of portal venous pressure.
After 10 seconds of stimulation the arterial pressure and heart rate begin to rise (*vagus escape*). (From an experiment by Dr. I. Calma.)

affected; the volume of the organs is therefore diminished. [If the heart is depressed for a *long* period, venous engorgement will work back to the periphery, causing swelling of the organs; this accounts for the enlarged congested viscera of chronic so-called congestive heart failure as seen clinically.] On ceasing to stimulate the vagus nerve, the blood pressure mounts rapidly and may exceed the initial pressure [1] before finally returning to normal.

[1] The after-rise is due partly to asphyxial stimulation of the vasomotor centre increasing arteriolar tone, and the resumption of vigorous activity by the distended heart.

(vi) "*Vagus Escape.*"—Sometimes, although vagal stimulation is continued, the heart begins to quicken, and the blood pressure rises. Several factors are possibly at work:

(*a*) The onset of vagus escape is often associated with high pressure in the *great veins* entering the right auricle (Fig. 155). This stimulates afferent fibres which pass up the (intact) vagus to the cardiac centre reflexly to accelerate the heart (auricular reflex, p. 272). This reflex *competes* with the effects of the artificial stimulation of the other (cut) vagus and may overcome it, causing the heart to increase in rate.[1]

(*b*) The lowered arterial blood pressure tends to accelerate the heart reflexly via the sino-aortic nerves (p. 272).

(*c*) In some cases complete heart block is produced, and the ventricles assume their independent rate, which is in no way affected by the vagus nerve.

Atropine annuls the effects of peripheral vagal stimulation on the heart (cf. p. 720).

2. **The Sympathetic.**—Anatomical details are given on p. 709. The sympathetic exerts an unimportant tonic accelerator action on the human heart. Bilateral excision of the stellate ganglia in man (which cuts off most of the sympathetic supply of the heart) has little effect on its rate at rest. Stimulation of the sympathetic nerves increases the *rate* and *force* of contraction, *i.e.* it accelelerates and augments the heart beat; it also enhances the excitability and irritability of the myocardium and may give rise to *ectopic beats* (*extrasystoles*) (cf. p. 286).

Regulation of the Heart Rate.[2]—There is believed to be a *cardiac centre* in the floor of the fourth ventricle in the region of the vagus nucleus. It is described as consisting of two (somewhat hypothetical) portions:

(1) *Cardio-inhibitory* centre, which presumably is part of, or closely related to, the dorsal nucleus of the vagus.

(2) *Cardio-accelerator* centre, in the same region, which has connections with the thoracic cord and so controls the sympathetic supply of the heart.

The rate of the heart is controlled mainly by *varying the degree of vagus tone* (*i.e.* vagal inhibitory activity). Unless a specific statement is made to the contrary, all the reflex and chemical adjustments of heart rate to be described are due *principally to changes in vagus tone and only to a minor degree to modifications of sympathetic activity.* If the vagus in man is completely "blocked" by atropine, the heart may accelerate to about 150 per minute. The further quickening which takes place in violent exercise—to about 180 per minute—is probably due to additional active stimulation by the sympathetic.

Vagus tone is minimal in infants in whom the heart rate may range round 120. Vagus tone increases with advancing age. In the adult the average heart rate is around 70 per minute, but it may be as low as 40 or 50 or reach 80 or 90.

As explained on p. 744, vagal tone is *reflexly* produced by afferent impulses in the sino-aortic nerves. Bilateral section of these nerves completely

[1] McDowall, *J. Physiol.*, 1926, *61*, 131.

[2] It is very difficult (and undesirable) to consider the control of the heart rate without at the same time discussing the regulation of the blood pressure. It is advisable to read pp. 309 *et seq.* in conjunction with this section.

abolishes vagal tone and produces the same degree of cardiac acceleration as bilateral vagotomy.

The rate of the heart is affected by—

1. Impulses from the higher centres.
2. Respiration.
3. Reflexes (1) from the aortic arch and carotid sinus and the carotid and aortic bodies.
   (2) from the great veins (" auricular " reflex).
   (3) from other parts.
4. O_2 and CO_2 content of blood.
5. Body temperature.
6. Intracranial pressure.
7. Adrenaline, thyroxine and vasopressin (pitressin).
8. Muscular exercise.

1. **Higher Centres.**—Emotion may readily affect the pulse rate: in states of excitement, the heart usually quickens; sudden shocks may slow or even arrest the heart. The common form of *fainting* attack provoked by emotion or long standing is partly due to vagal overaction; nervous women, convalescents, and patients with aortic incompetence are particularly liable to be affected. The attack is ushered in by unsteadiness, giddiness, and dimness of vision; the heart's action gradually becomes slower, and the pulse rate may fall to 50 or 40. At the height of the attack, pallor and sweating may be prominent symptoms. The systolic blood pressure may sink to 50 mm. Hg, owing to *independent depression of the vasomotor centre*; the entire syndrome is sometimes called a *vaso-vagal* attack.[1]

The influence of the *cerebral cortex* and *hypothalamus* on heart rate is referred to on pp. 671, 716.

2. **Respiration.**—In most adults, during quiet breathing, the heart rate does not alter. During voluntary deep breathing, all normal people show quickening of the heart with inspiration, and slowing with expiration. This respiratory variation occurs in many healthy children with quiet breathing. The respiratory variation in the heart rate is referred to as *sinus arrhythmia.* It has no pathological import whatever.[2]

Sinus arrhythmia is due to alterations in vagus tone which take place with each phase of the breathing; it is abolished by atropine or vagal section. It has been attributed (i) to afferent impulses from the *lungs*, (ii) to variations in pressure on the venous side of the *heart,* or (iii) to some *central* influence. If the isolated head, connected with the trunk by means of the vagi only, is artificially perfused, it is found that the arrhythmia of the heart does not follow the rate of the artificial respiration which is applied to the trunk, but corresponds to the discharge rate of the *respiratory centre* as determined by recording the movements of the alæ nasi (Fig. 156); this correlation becomes even more evident when the respiratory movements in the head are stimulated. The arrhythmia persists after denervation of the lungs and is independent

[1] Profound slowing of the heart occurs in association with the faint produced by a severe hæmorrhage (cf. p. 84).

[2] *Irregular Variations.*—Periodic variations in the rate of the heart may occur independently of respiration and without apparent reason. They may be noticed in convalescence from acute fevers and after the administration of digitalis. (Lewis.)

of the fluctuations in the pressure of the blood which is perfused through the brain. Heymans concludes from these observations that the arrhythmia is to an important degree due to direct *irradiation of impulses from the respiratory to the cardiac centre.*

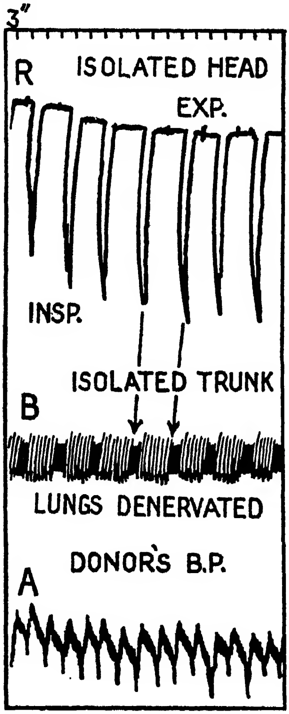


FIG. 156.—Analysis of Sinus Arrhythmia.

R = Record of respiration taken from alæ nasi. The head has been isolated from the trunk by division of the neck, the vagi alone being left intact. The head is kept alive by perfusion from a donor.
A = Blood pressure of donor.
B = Record of heart rate of isolated trunk; this is kept alive by means of artificial respiration. Note the marked sinus arrhythmia. The black patches in the record are regions where the rate is too rapid for the individual beats to appear. The arrhythmia is obviously related to the *rhythm of respiration in the head.* The fast periods correspond to inspiration and the slow to expiration.
Time in 3 seconds. (Heymans, *Ergeb. Physiol.*)

3. (1) **Sinus and Aortic Nerves.**—Experiments fully described on pp. 738 *et seq.*, prove that the *sinus* and *aortic nerves stabilize the heart rate* in the normal resting individual, are responsible for resting vagal tone and reflexly modify heart rate in response to sudden changes in arterial blood pressure. The principal points are the following:

(i) Stimulation of the central end of the aortic or sinus nerves usually reflexly slows the heart.

(ii) Section of these nerves accelerates the heart to the same degree as injection of atropine or bilateral vagal section, *i.e.* vagal tone is completely abolished and cardiac irregularities are set up (p. 744).

(iii) In perfusion experiments a rise of pressure in the carotid sinus reflexly slows the heart, while a fall of pressure reflexly quickens the heart.

The carotid sinus endings respond most actively to changes in blood pressure when the initial pressure is about the normal for the species. This is well shown in Fig. 157. A fall of pressure from 125 mm. to 100 mm., or a rise from 125 mm. to 150 mm. produces far greater reflex quickening or slowing respectively than do changes of equal magnitude farther up or down the blood pressure scale; the aortic nerve endings act similarly.

These observations account for the inverse relationship commonly noted in the *resting* animal between blood pressure and heart rate (Marey's law); *i.e.* when the blood pressure rises, the heart rate falls, and vice versa. It should be noted, however, that in *exercise, emotion, anoxia,* CO_2 *excess,* and other conditions, *hypertension may be associated with a rapid heart rate.*

The rapid resting heart rate in children is perhaps related to their lower arterial blood pressure level.

(2) **Afferents from Great Veins.**—Nerve fibres arise from the *venous* side of the heart (chiefly from the root of the great veins) (p. 735) and pass in the vagus trunk to the cardiac centre. These nerve endings are stimulated when the venous pressure rises. As a result vagus tone is reflexly depressed

and the heart rate rises; the sympathetic accelerator fibres are also reflexly stimulated to a minor extent. This reflex is called the *auricular* or *Bainbridge* reflex (strictly speaking is should be called the *venous* reflex). It can be demonstrated by injecting a large volume of saline rapidly into the central end of the jugular vein; if the vagi are intact the heart accelerates. In *muscular exercise* the venous return is greatly increased; the resulting rise of venous pressure plays an important part in bringing about cardiac acceleration during muscular exercise, thus helping the heart to get rid of all the blood reaching it.

In *heart failure,* engorgement of the right auricle often occurs, associated with a general rise of venous pressure (p. 319). The venous reflex is set up, and accounts in part for the rapid pulse rate found in this condition. The cardiac acceleration which occurs in the early stages of *hæmorrhage* in spite

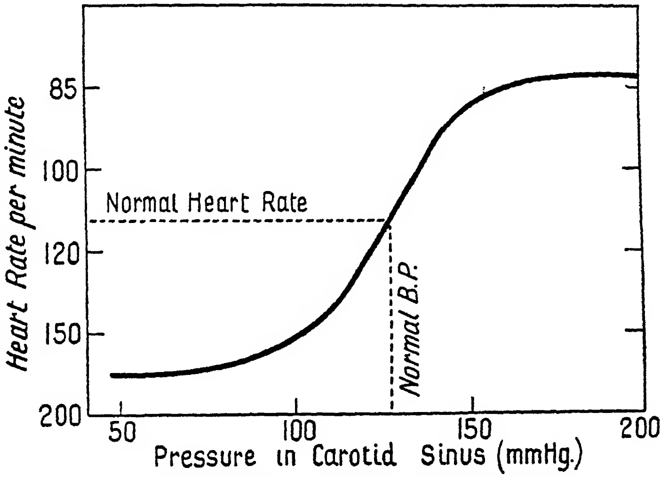


FIG. 157.—Carotid Sinus Pressure and Heart Rate.

The curve shows the effects on heart rate of increasing the pressure in the isolated carotid sinus. At the pressure is raised the heart rate is decreased. Normal carotid sinus pressure produces a hears rate of normal frequency. (Koch, *Kreislaufforsch.*, 1930.)

of low venous pressure, is reflexly produced by anoxia (stimulating the chemoreceptors) and the fall of *arterial* blood pressure (stimulating the pressure receptors).

A rise of central venous pressure also reflexly accelerates *respiration*; this may be a factor in the production of cardiac dyspnœa (p. 461).

(3) OTHER AFFERENTS.—Stimulation of *sensory* nerves may produce varying effects. Stimulation of the central end of the fifth nerve or of the splanchnic nerve causes slowing of the heart. A blow on the abdomen or irritation of the mucous membrane of the nose produces similar results. Most painful stimuli quicken the heart.

4. **Changes in Blood Oxygen and CO_2.**—(i) Lack of oxygen—*anoxia*—accelerates the heart (Fig. 182); this is another factor in the production of the rapid pulse of heart failure. The degree of acceleration is related to the severity of the anoxia. When an inert gas like nitrogen is inhaled to

produce acute and severe anoxia the pulse may increase in 1–2 minutes to 140 per minute. An increased pulse rate is likewise found in anæmic states, carbon monoxide poisoning, during the early stages of hæmorrhage (p. 83) and at high altitudes.

Anoxia first causes an apparent increase in the force of the heart beat, but the vigour later dies down, though the heart remains accelerated till just before the end (cf. Fig. 243).

(ii) *Excess of CO_2* when of moderate extent also quickens the heart; this can be readily seen in man when 5% CO_2 is breathed for several minutes to produce CO_2 accumulation in the alveoli and the arterial blood (Fig. 181). In *very large amounts*, CO_2 poisons the bundle of His, produces heart block and a slow ventricular rate; this is seen in the terminal phase of asphyxia.

The cardiac acceleration produced by anoxia and CO_2 excess is due partly to a direct action on the cardiac centre and partly to *reflex stimulation via the chemoreceptors* in the carotid and aortic bodies (p. 745).

5. Temperature.—A rise of *body temperature* acts (i) directly on the sino-auricular node and quickens the rate of impulse formation (p. 236); (ii) the cardio-accelerator centre is also stimulated (cf. pp. 475 477.)

6. A rise of *intracranial pressure*, if very large, such as occurs *late* in cases of cerebral tumour, directly stimulates the vagus centre and slows the heart.

7. (i) Adrenaline.—See p. 725.

(ii) Thyroxine accelerates the heart in part by a direct action on the pace-maker. The heart is markedly quickened in hyperthyroidism (p. 990) and is slower than normal in myxœdema (p. 983).

(iii) Vasopressin (Pitressin).—See p. 46.

8. Muscular Exercise.—The cardiac acceleration of exercise is due, as stated, mainly to the venous ("auricular") reflex. Accessory factors are impulses from the higher centres, the rise of CO_2 tension and temperature (and possibly anoxia and adrenaline secretion (cf. p. 433)).

THE OUTPUT OF THE HEART[1]

The terms "minute output" or "circulation rate" mean output per ventricle per minute; the term "stroke volume" means output per ventricle per beat.

Control of Cardiac Output.—The output of both ventricles under *normal* conditions is exactly the same. The volume of blood pumped out by the left ventricle depends on the amount it receives from the right side of the heart via the pulmonary circulation. The output of the heart depends on (i) the *venous return*; (ii) the *force* and (iii) *frequency* of the beat. It is also intimately related to (iv) *blood pressure*.

1. **Venous Return.**—This depends chiefly on the following factors:

(i) The contractions of the *skeletal muscles* squeeze the veins contained within them and drive the blood on towards the heart. The veins possess valves which prevent regurgitation of the blood in the intervals between the muscular contractions (the veins are also actively contractile) (cf. vii, *infra*.)

[1] Grollman, *Cardiac Output of Man in Health and Disease*, Springfield, Ill., 1932.

(ii) With each inspiration there is an increased *negative pressure* in the pleural cavity which aspirates blood towards the heart (cf. p. 368). The diaphragm descends, raises the intra-abdominal pressure, and so squeezes blood out of the abdominal veins.

(iii) There is a slight *positive pressure* in the capillaries and a somewhat lower pressure in the veins which can propel blood on towards the heart. If the general arterial pressure does not fall, relaxation of the arterioles supplying any region raises the capillary pressure locally, and so aids the venous return (*e.g.* from the active muscles in muscular exercise).

(iv) *Gravity* assists the venous return from parts above heart level, but greatly hinders the return from the dependent regions.

(v) *Changes in blood volume*: the venous return is diminished by a decrease in blood volume as after hæmorrhage (Fig. 46 and p. 82); it is correspondingly increased after intravenous injection of saline (p. 59).

(vi) *Capillary tone:* normally many of the capillaries in any organ are quite narrow and some may be closed. If all the capillaries over a considerable area become patent and widely dilated, they can accommodate a large proportion of the blood volume, and less is available to return to the veins. The venous return is greatly reduced for this reason in histamine poisoning (p. 337).

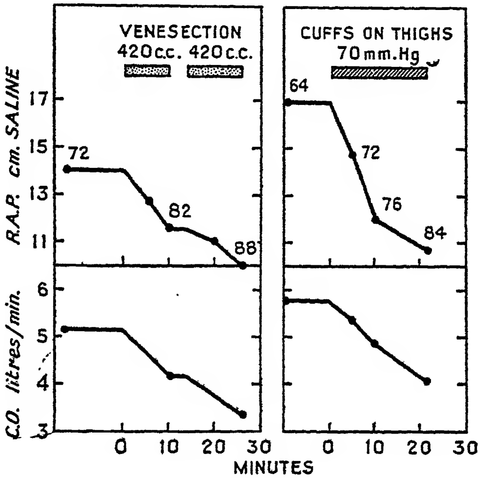


FIG. 158.—Results of Venesection in Man.

FIG. 159.—Results of "Physiological Venesection" in Man.

Normal subject in supine position.
R.A.P.=right auricular pressure (in cm. saline anterior to posterior surface of thorax) recorded by catheter in right auricle connected to manometer.
C.O.=cardiac output (in litres per minute) recorded by direct Fick method.
Figures on R.A.P. curves=heart rate per minute.
In Fig. 158, withdraw 840 c.c. of blood in 25 minutes.
In Fig. 159, apply cuffs to both thighs and inflate to diastolic blood pressure (70 mm. Hg) to trap approximately 700 c.c. of blood in legs.
Note similarity of results in both experiments. (McMichael and Sharpey-Schafer, *Brit. Heart J.*, 1944, *6*, 33.)

(vii) The veins themselves are controlled by the vasomotor centre (p. 318) and by chemical influences in very much the same way as the arterioles. Any factor which constricts the veins and so decreases their capacity helps to increase the return of blood to the heart.

Other things being equal, an increase in venous return increases the cardiac output per minute and a decrease in venous return correspondingly decreases the cardiac output (Figs. 158, 159 and 46).

2. Force of the Heart.—This depends on:

(i) *The Initial Length of the Muscle Fibres.*—Within limits, the greater the initial length of the heart muscle fibres at the beginning of systole, the more forcible is the contraction ("law of the heart" (Starling)). This

can be well shown experimentally in the denervated heart-lung preparation, in which the heart rate is constant.[1] It is found that if the *venous return*

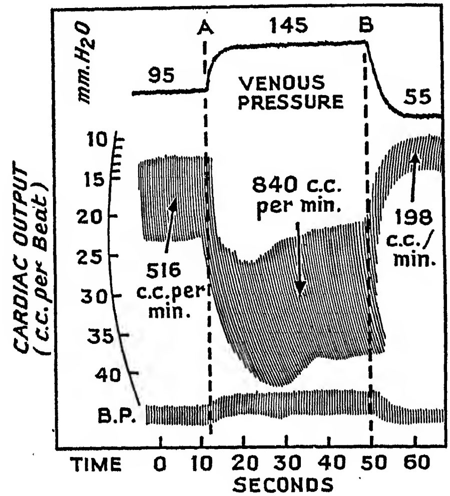


FIG. 160.—Effect of Increasing Venous Pressure and Venous Inflow on Ventricular Outpu per Beat in Heart-Lung Preparation. (Starling *et al.*, *J. Physiol.*, 1914, *48*.)

Records from above downwards:
- Venous pressure in mm. H_2O.
- Cardiac output per beat recorded with ventricles in cardiometer. The scale on the left shows the output *per beat* (per ventricle) in c.c.; the values on the record give the output (per ventricle) in c.c. per minute. Upstroke=systole=*decrease* in ventricular volume; downstroke=diastole=*increase* in ventricular volume. As the cardiac nerves were cut no change in heart rate occurs.
- BP: arterial blood pressure.
- Time in seconds.
- At A (first vertical) the venous pressure was increased from 95 to 145 mm. H_2O and the venous inflow was correspondingly increased. The ventricles undergo progressive dilatation, *i.e.* they fail initially to discharge as much blood as they receive. The increased stretch of the muscle fibres during diastole leads to increased force of contraction; when compensation is fully established the cardiac output is increased from 516 to 840 c.c. per minute. Note, however, that the volume at the end of systole remains increased, *i.e.* the ventricles are not now emptying during systole as completely as during the control period.
- [It is probable that under more normal experimental conditions, although the ventricles would dilate during diastole as shown in this experiment, they would empty more completely during systole.]
- At B (second vertical) the venous return was lowered to 55 mm. H_2O. The diastolic volume of the heart markedly decreased and the cardiac output fell to 198 c.c. per minute. The arterial blood pressure changed very little.

is artificially increased with resulting distension of the heart, the *force of the beat is correspondingly enhanced,* and the larger inflow is successfully dealt with

[1] This preparation is useful for studying the output of the heart experimentally under conditions where some of the factors can be modified at will. The thorax of an animal is opened and artificial respiration is carried out. The aortic arch is tied beyond the origin of the innominate artery. By means of a cannula the blood is led from the innominate artery to an artificial peripheral resistance which consists of a thin rubber tube enclosed in a glass tube, in which the pressure can be raised to any desired extent. The blood is warmed and returned to the right auricle and sent by the right ventricle through the lungs and thus back to the left heart. Variations of pulse rate are prevented by section of the vagi. The output of the heart can be measured directly by collecting the blood for a given time after it has passed through the peripheral resistance.

by a larger cardiac output without change in heart rate (see Fig. 160 and legend).

In accordance with this principle, it is found that when the ventricle is underfilled, *e.g.* after hæmorrhage, or in obstruction of the mitral orifice, the force of the beat is reduced; when the venous inflow is large and diastolic filling is great, as in muscular exercise, powerful beats are obtained. But should the fibres be *overstretched* because of excessive filling, a more feeble contraction is obtained, the output per beat is diminished and the heart does not empty itself effectively. This may occur in certain stages of clinical cardiac failure and is then associated with engorgement of the venous side of the heart. In such cases, venesection, by lowering the venous pressure and reducing the abnormal degree of stretch of the muscle fibres, may greatly improve the efficiency of the individual beats and consequently of the state of the circulation.

Chronic excessive stretching of a mild grade may serve as a *growth stimulus* to heart muscle fibres. This is well seen in certain instances of valvular incompetence. Thus, in mitral or aortic regurgitation, there is excessive accumulation of blood behind the leak, *i.e.* in the left auricle and left ventricle respectively. The fibres respond to stretch by more powerful contraction, and in the course of time overgrowth (*hypertrophy*) of the chamber develops (p. 297).

(ii) *Diastolic Pause.*—If the diastolic pause is too short, the heart has inadequate time to recover from the effects of the previous contraction, and the force is diminished. In addition, with a constant rate of venous return, a shorter diastole may mean lessened venous filling and consequently feebler beats (cf. *infra*).

(iii) *Nutrition and Oxygen Supply of the Heart.*—Obviously these must be adequately maintained; the subject is discussed on pp. 236 *et seq.*

3. **Frequency of the Heart.**—The rate of the heart influences the cardiac output, both directly and by affecting the force and the venous filling. In the resting person, with *constant venous return*, the normal range is satisfactory, because it provides ample diastolic time for venous filling and for recovery of the heart muscle, ~~and the~~ beats are frequent enough for effective transfer of the blood from the venous to the arterial side. *Marked* alterations in the frequency (with unaltered venous return) serve no useful purpose, and may be definitely harmful. This is well seen in cases of *paroxysmal tachycardia* and *heart block*. In the former condition the heart may accelerate to 150 or 200 per minute. Diastole is too brief to allow proper filling of the heart to take place, and each beat is very small; in spite of the high frequency of the heart, the output *per minute* may be reduced to less than half the normal (cf. p. 290). In *heart block*, on the contrary, the ventricular rate is very slow. There is ample time for venous filling, and for recovery processes to occur. But the beats may be too infrequent to empty the venous side, and consequently, though each beat is large, the output per minute is reduced (p. 284).

If the *venous return is increased*, the heart can cope with it, within limits, by increasing the size of each beat. But this adjustment will only provide for a two or threefold increase in output in man, corresponding to an increase in the stroke from say 60 to 180 c.c. per beat. When the venous inflow is still greater, as in severe muscular exercise, acceleration is essential if the heart is to dispose

venous blood, which is only found in the right heart and the pulmonary artery.

The modified methods which may be used in man are described below.

(1) **Direct Fick Method in Man using Oxygen.**[1]—(i) The *oxygen consumption* is determined in the usual way (p. 372).

(ii) *Arterial Blood.*—As it is undesirable to puncture the radial artery the O_2 content of the arterial blood is calculated from a hæmoglobin determination (p. 175), assuming that the blood has the normal average degree of O_2 saturation (*i.e.* it is 95% saturated).[2]

(iii) *Mixed Venous Blood.*—To obtain a sample of mixed venous blood, a long, fine rubber tube is introduced with suitable precautions into the right antecubital vein and is directed up along the veins until it *enters the right auricle* (Cournand's method). The patient's interest and safety should always be conscientiously considered before this procedure is embarked upon. The oxygen content of the blood so collected is determined directly and the cardiac output calculated.[3] The technique involves an inevitable degree of emotional strain on the part of the subject and so gives readings "at rest" which are higher than the true basal values. On the other hand it provides reliable information about *changes* in the cardiac output which take place under different experimental and clinical conditions.

(2) **Indirect Fick Principle (using Oxygen) applied to Muscular Exercise (Hill).**—The Fick method in a modified form can give information about the cardiac output even in violent exercise. The *arterial oxygen content* and the *O_2 consumption* are determined as in (1) *supra*; in severe exercise the latter may reach 4 litres per minute or more. The O_2 content of the *mixed venous blood* is approximately assessed from the following considerations and not directly determined. During bicycle work, when the O_2 consumption is about 2 litres per minute, the O_2 content of the mixed venous blood may be reduced by 60%, *i.e.* to about 7 c.c.%. In more strenuous work it is justifiable to presume that the blood is even more extensively reduced, *i.e.* the mixed venous O_2 content is perhaps 4 c.c.%. We know that the blood leaving the muscles is almost completely reduced, but that from the skin, brain, and viscera still contains fair amounts of oxygen. The usual calculation is employed:

$$\text{Cardiac output} = \frac{O_2 \text{ consumption (c.c.)}}{\text{Arterial-venous } O_2 \text{ difference (c.c.\%)}} \times 100$$
$$= \frac{4000 \times 100}{19-4} = \underline{27 \text{ litres.}}$$

This value is undoubtedly reliable as far as order of magnitude of the cardiac output is concerned.

[1] Cournand *et al.*, *Proc. Soc. exp. Biol. Med.*, 1941, *46*, 42; *Surgery*, 1943, *13*, 964; *J. clin. Investig.*, 1942, *21*, 287; 1945, *24*, 106; McMichael and Sharpey-Schafer, *Brit. Heart J.*, 1944, *6*, 33.

[2] *Example.*—Hæmoglobin content=15 g-%; 1 g. of hæmoglobin unites with 1·34 c.c. of O_2 when fully saturated; ∴ if the blood is 95% saturated, the oxygen content is:

$$15 \times 1{\cdot}34 \times \frac{95}{100} = 19 \text{ c.c.\%}$$

[3] Even in the right auricle the blood from the inferior and superior vena cava may be partially streamlined so that a true *mixed* venous sample is not always obtained.

The *minimal* cardiac output that would enable the O_2 consumed to be absorbed in the lungs can be calculated by assuming the mixed venous O_2 content to be zero. In the above example the cardiac output would be $\frac{4000\times100}{19}=21$ litres. As the mixed venous O_2 content is undoubtedly more than zero, the cardiac output in this form of exercise *must* exceed 21 litres per minute.

If the O_2 consumption is 5 litres, the cardiac output may be calculated to be about 35 litres. If the pulse rate is 180 per minute, the stroke volume is 150–200 cc.[1]

Factors Influencing Cardiac Output in Man.—(1) Resting Output.— With the direct Fick method (using oxygen) the cardiac output in adults is 5–6 litres per minute; as these conditions are not basal owing to the presence of emotional tension, the true resting cardiac output may be about 4–5 litres. The output per beat is about 60 or 70 c.c. There is a fair relationship between cardiac output and body weight and a closer relationship between output and *surface area* (p. 378). Little is known about the output in children or in old age. *Emotional states* produce varying degrees of elevation of cardiac output.

In passing *passively* from the horizontal to the erect position the cardiac output diminishes by about 1 litre per minute. Under these experimental conditions the leg muscles are slack and are not pumping the blood back to the heart; as gravity hampers the venous return the cardiac output falls. Under conditions of *normal standing* the erect position is accompanied by some degree of limb-muscle activity which adequately compensates for the deleterious effects of gravity.

(2) Distribution of Cardiac Output.—Partial data (given in c.c. *per minute*) are now available for man about the distribution of the cardiac output to the various organs at rest and during activity. The outstanding finding is that the two *kidneys* receive about 1300 c.c. of the total resting output of 5000 c.c. (p. 36). The blood flow to the *brain* is estimated at about 750 c.c. (p. 306). The blood flow to resting limb *skeletal muscle* varies between 1 and 4 c.c. per 100 c.c. of muscle (p. 308). The muscles form 40% of the body weight and in a 70-kg. man their total weight is 28 kg. If limb blood flow conditions are representative of the musculature as a whole the total resting muscle blood flow is 300–1100 c.c. At a guess, the *coronary* artery flow is 100 c.c. The kidneys, brain, skeletal muscles, and heart thus account for 2500–3000 c.c. The other half of the resting cardiac output passes to the skin, abdominal viscera (excluding the kidneys), and bones.

(i) The blood supply to the *skin* is mainly adjusted to meet the needs of *temperature regulation.* The degree of possible change of blood vessel calibre varies with the skin region examined and is greatest in the hands (especially the fingers), the feet, and probably the face. The flow (*per* 100 *c.c. volume*) in the two terminal finger phalanges may vary from 0·2 c.c. under conditions of extreme vasoconstriction to 120 c.c. in full vasodilatation; in the fingers as a whole the maximum flow is 60 c.c. (5% of the finger is skin) (Fig. 198).

(ii) The blood flow through *muscle* varies with its *metabolic activity.* In vigorous activity the forearm flow (per 100 c.c.) may rise to 30 c.c. (Fig. 262). If such a level of muscle blood flow became general the total muscle flow

[1] For ballistocardiographic method, see Starr, *Harvey Lect.*, 1947, *42*, 194.

would amount to about 14,000 c.c. Such an order of flow necessitates, of course, a great increase in the total cardiac output.

(iii) The *coronary* flow during violent activity rises to 800 c.c. (p. 238).

(iv) The *brain* blood flow is probably unaffected by the level of mental activity.

(v) The flow in the *abdominal viscera* increases considerably during secretion of the digestive juices (Fig. 519).

(3) **Relation of Output to Metabolism.**—(i) The metabolic activities of the body are limited by the oxygen supplies. Each litre of O_2 employed in the combustion of foodstuffs makes available 4·8 kg.-calories (on an average). The output of the heart, and consequently the oxygen supplied to the tissues, are related (as might be expected) fairly closely to the *metabolic rate*. Bodily activity means more vigorously contracting muscles, more powerful respiratory movements, relaxed arterioles in muscles, and raised capillary pressure. The heart deals with the increased venous return in mild exercise either (*a*) by increasing the *rate* or (*b*) by increasing the *stroke volume*, or (*c*) both methods may be employed. In addition, the increased demands of the tissues for oxygen are met by *abstracting more O_2 from the blood reaching the tissues, i.e.* the coefficient of utilization (arterial-venous oxygen difference) is raised, and the venous blood leaves more extensively reduced (cf. p. 413).

(ii) In *strenuous exertion*, such as severe running, the minute output may reach 30–40 litres; the pulse rate rises to 150–180 per minute, and the stroke volume may be as large as 170–220 c.c. The mixed venous blood is extensively reduced, *e.g.* the oxygen content may fall to about 4 c.c., and the arterial-venous oxygen difference may be, *e.g.* 19—4 c.c.=15 c.c.%. If the blood volume is taken to be 5 litres, it is obvious that, at rest, an amount little less than the blood volume is discharged each minute from each ventricle. In strenuous exertion however up to six to eight times the blood volume (about 8 gallons) may pass through each ventricle per minute. This last figure probably represents the maximum possible effort of the heart (cf. p. 433).

(iii) As indicated on pp. 990 and 983, owing to the changes in metabolism in *exophthalmic goitre* and *myxœdema*, corresponding changes in cardiac output are encountered. It is somewhat difficult to see how alterations in the metabolism (not the result of exercise) produce the appropriate modifications in the venous return which must of necessity precede an alteration in cardiac output. It may be supposed, however, that when the metabolism is increased, the arterioles in the more active tissues are relaxed, with a corresponding rise of capillary and venous pressure which propels blood more rapidly to the heart.

(iv) The ingestion of *food* (which stimulates metabolism, increases the oxygen consumption and raises the pulse rate) may increase the output per minute by 0·5–1 litre. Changes in output of a similar order of magnitude are said to follow the ingestion of large volumes of *fluid*.

(4) **External Temperature.**—(i) When the external temperature is *lowered* (from 16°–20° C. as the starting-point), the minute output and pulse rate are little altered. The metabolism is increased, but the greater oxygen requirements of the tissues are provided solely by more extensive reduction of the blood.

(ii) If the external temperature is *raised* the metabolism does not change at first. Even when the external temperature is as high as 40°–45° C. the

oxygen consumption is increased very slightly; but both the pulse rate and cardiac output per minute are raised considerably. This increase in cardiac output is not due to the change in metabolic activity (which is small) but to the need for *temperature regulation*. External heat causes dilatation of the arterioles of the skin (and also of the capillaries); this enables a much larger volume of blood to reach the skin and facilitates the giving off of heat particularly by evaporation and radiation. The blood in the skin veins has a high oxygen content, which raises the oxygen content of the mixed venous blood. The increase in venous return needed to maintain the raised minute output is presumably brought about by raised capillary and venous pressure resulting from the cutaneous arteriolar dilatation (cf. p. 477).

(5) Relationship to Body Weight.—If weight reduction is brought about by dietetic means in *obese* subjects both the cardiac output and the work of the heart are substantially reduced (*e.g.* by 30%).

(6) *Action of Anoxia.*—At high altitudes the cardiac output slowly rises at first and then decreases once more to normal (cf. Fig. 243, p. 399).

(7) *Action of CO_2 Excess.*—The inhalation of carbon dioxide-rich mixtures is said to be without effect on the cardiac output until the concentration of the gas in the inspired air exceeds 6%. With higher concentrations some increase in cardiac output is observed (cf. Fig. 181).

For effects of *hæmorrhage* see p. 81; injections of *intravenous saline,* see p. 59.

ABNORMALITIES OF CARDIAC RHYTHM

Heart Block.[1]—By heart block is meant a condition in which there is defective conduction in some part of the heart.

The following varieties of heart block may be recognized:

(1) **Sino-Auricular Heart Block.**—This is a rare condition in which a whole heart beat is lost at varying intervals. During this pause there is neither auricular nor ventricular contraction, nor does any electrical variation occur. After an interval which is usually less than two complete cycles, the heart resumes its normal action. Occasionally every other beat is suppressed completely, and a profound slowing of the whole heart to about 30 beats per minute is noted. The condition is at once unmasked by exercise, when the "missing" beats reappear, and the heart suddenly doubles its rate. It then gradually accelerates further, like a normal heart.

S.A. heart block may be produced by vagal stimulation, and is relieved by "paralysis" of the vagal endings with atropine, or by exercise (which decreases vagus tone). The clinical condition is perhaps due to heightened vagus tone acting on a susceptible S.A. node. Block is produced *within* the substance of the node, and the impulse never gets out to activate the rest of the heart.

(2) **Main Bundle Block.**—If the bundle of His is compressed or damaged, conduction is impaired; changes occur in the following order:

(i) Delayed conduction (P–R interval exceeds 0·2 second).

(ii) Failure of occasional and then of a larger proportion of the auricular impulses to reach the ventricles (see Fig. 161).

(iii) The ventricles respond to every second, third, or fourth impulse from

[1] Rosenman *et al.*, *Arch. int. Med.*, 1950, *86*, 196.

the S.A. node, as shown by the establishment of ratios like 2 : 1, 3 : 1, 4 : 1 between the auricular and ventricular beats.

The above conditions are called *partial* heart block.

(iv) *Complete* heart block: none of the auricular impulses reach the

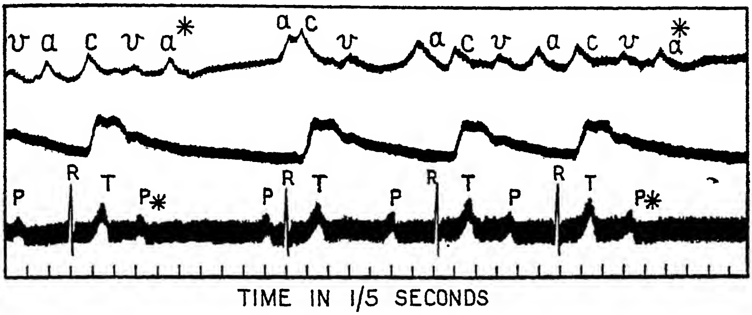


FIG. 161.—Simultaneous Venous, Radial, and Electrocardiographic Curves from Patient with Partial Heart Block, showing " Dropped " Beats.

Records from above downwards: jugular venous pulse, radial pulse, and electrocardiogram, Lead II. The second P wave (P*) in the electrocardiogram is not followed by a ventricular complex; similarly the second *a* wave (*a**) is not followed by *c* and *v* waves; nor is there any corresponding pulsation at the wrist. Following this dropped beat (long pause in the pulse tracing) the P-R and *a-c* intervals are normal; in the succeeding cycles they lengthen to 0·4 and 0·5 sec. Notice that the *increase of the second P-R (and a-c) intervals over the first is greater than the increase of the third over the second*; this shortens the interventricular period directly preceding the next dropped beat, *i.e.* the ventricle quickens to the point of disturbance. Following the next dropped beat the events are repeated. Time (lowest record) in fifths of a second. [Cf. Fig. 476.] (Lewis, *Mechanism and Graphic Registration of Heart Beat.*)

ventricles. The beats of auricles and ventricles are completely dissociated, and bear no relation whatever to one another. The ventricles now beat with an independent rhythm (Fig. 162).

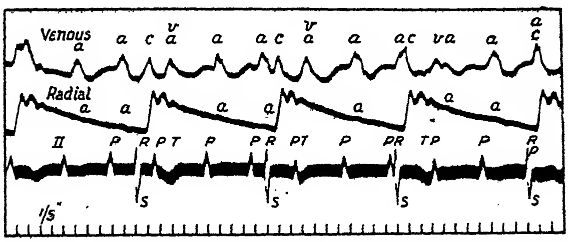


FIG. 162.—Complete Heart Block in Man. (Lewis, *Mechanism and Graphic Registration of Heart Beat.*)

Records from above downwards are: jugular venous pulse, radial pulse, and electrocardiogram (Lead II). Note the regular relationship between the jugular *c* wave, the radial primary wave, and the R wave (all indicative of ventricular activity). There is a regular relationship also between the jugular *a* waves and the P wave (indicating auricular activity). In their respective records the *a* waves and P waves follow regularly. There is no relationship between auricular and ventricular events.

When 2 : 1 block is established, the electrocardiogram shows regularly spaced P waves, which are twice as numerous as the ventricular complexes. When complete block is present, the independent ventricular beats arise from the most rhythmic part of that chamber which is usually the region

of the bundle just below the site of the block. The excitation process, therefore, reaches the two ventricles along the normal channel of the two branches of the bundle. The ventricular complex may be quite normal in character, but is in no way related in time to the auricular wave.

EFFECTS OF HEART BLOCK.—While the heart block is partial, the ventricles beat slowly but always in response to auricular impulses. The circulation is fairly effectively carried out at rest, the diminished rate being compensated for by the larger output per beat. Because of the larger output per beat the systolic pressure tends to be raised ; the diastolic pressure is lower than normal owing to the longer time available for the blood to leak out during diastole from the arterial system into the capillaries.

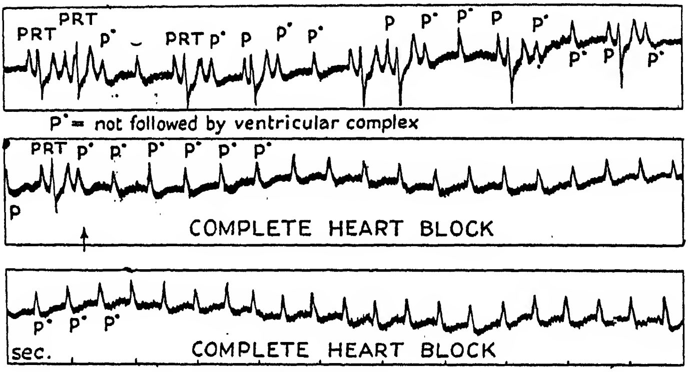


FIG. 163.—Sequence of Events in Stokes-Adams' Attack in Man.

Electrocardiograph records to be read from left to right and successively from above downwards. Time in sec.

In upper record there are frequent dropped beats (P· not followed by ventricular complex) interspersed between complete PQRST complexes.

At arrow in middle record, heart block becomes complete and from then onwards there is a regular succession of P· waves, none of which is followed by ventricular complexes. (Hermann *et al.*, *Arch. Malad. Cœur.*, 1937, *30*, 772.)

Exercise is greatly limited because the necessary acceleration of the heart to cope with the greater venous return cannot take place.

If the heart block becomes *complete suddenly*, the ventricles may stop beating temporarily. [The same result is obtained when a ligature is tied between the auricle and ventricle of the frog ("Stannius ligature").] Some little time elapses before they can accustom themselves to the new conditions, and so the ventricles are quiescent until they begin beating at their own rate.

STOKES-ADAMS' SYNDROME.—Fig. 163 shows the electrocardiographic changes in a case in which, after a phase of partial heart block with frequent dropped beats, complete heart block developed acutely, leading to arrest of ventricular contraction for 2 minutes.[1] With ventricular standstill the blood pressure falls to zero and the blood supply to the brain ceases. After 5 seconds, consciousness is lost ; after 15 seconds there are muscular twitchings, convulsions and cyanosis ; breathing is intensely stimulated (presumably reflexly by the fall of blood pressure and the anoxia, and also by the CO_2 accumulation). After 30 seconds, breathing becomes slowed (owing to the

[1] Hermann *et al.*, *Arch. mal. Cœur*, 1937, *30*, 753.

direct depression of the respiratory centre by bulbar anæmia) and stops after 1 minute. If the heart remains quiescent death occurs after about 2 minutes. In the case illustrated by Fig. 163 the ventricles resumed beating; this led to a transient hypertension, no doubt due to vasoconstriction resulting from asphyxial stimulation of the vasomotor centre; later breathing returned.

An attack of complete heart block accompanied by fainting as just described constitutes the *Stokes-Adams' syndrome.*

(3) **Bundle Branch Lesions.**—If (for example) the right branch of the bundle of His is destroyed the left ventricle is invaded normally. In the case of the right ventricle, however, the impulse cannot reach it initially along the normal channels; the impulse spreads relatively slowly from the left to the right ventricle by muscular continuity until at some stage it

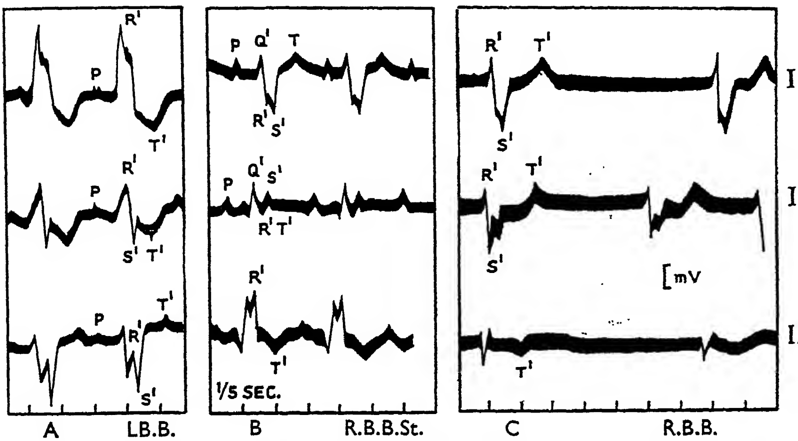


FIG. 164.—Electrocardiographic Appearances in Branch Bundle Lesions in Man.

A. Left branch bundle lesion.
B. Right branch bundle lesion ("standard" curve).
C. Right branch bundle lesion (common curve).

Leads I, II and III from above downwards. Time in 1/5 sec. (Modified from Evans and Turnbull, *Lancet*, 1937, ii, 1129.)

reaches the right branch bundle below the site of the block, after which invasion is completed rapidly. The pattern of left ventricular invasion is consequently abnormal and its duration is prolonged. The QRS complex which represents total ventricular invasion will likewise be prolonged (its duration always exceeds the normal upper limit of 0·12 second) and shows abnormal features; as the pattern of repolarization is also affected the ST segment and T wave will show abnormal features.

(i) LEFT BRANCH BUNDLE LESIONS.—The typical human findings are illustrated in Fig. 164 In all leads QRS exceeds 0·12 second. In Lead I the R is tall and is followed by a plateau which may show irregular variations ("splintering"). In Lead III there is a predominant downwardly directed and prolonged S. There is no isoelectric position of the ST segment. The T wave may take off from above or below the isoelectric level; its direction is in the opposite direction to that of the main initial deflection, *i.e.* downwards in Lead I and upwards in Lead III.

(ii) RIGHT BRANCH BUNDLE LESIONS.—The electrical findings are complex. In Lead I there is always a steep prolonged irregular downwardly directed S, followed by an upright T (Fig. 164, B, C). The findings in Lead III are however very variable; the main initial wave is commonly directed *downwards* (Fig. 164, C); in the so-called "standard" curve (which is *less* common) the main initial wave is directed upwards (Fig. 164, B).

New Rhythm Centres.—In the normal heart several centres of rhythm formation are potentially active, but the centre which generates an impulse most rapidly sets the pace of the heart. This is normally the S.A. node.

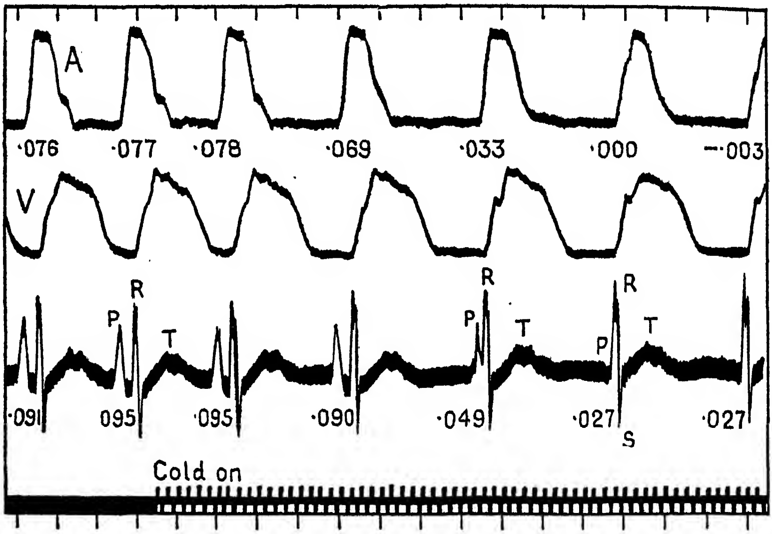


FIG. 165.—Myocardiographic Curves (A=Auricle; V=Ventricle) and Electrocardiogram from Lead II, showing the Effect of applying Cold to the Sulcus Terminalis in a Dog.

When cold is applied (see signal) the S.A. rhythm slows; at the fifth cycle there is escape of the A.V node, though the auricle, being already in systole, does not respond, *i.e.* the auricle responds to the impulse from the S.A. node, but the ventricle responds prematurely to the A.V. nodal impulse (the P–R interval is reduced to 0·049 second). Subsequent cycles show fully established A.V. rhythm. The auricle is represented in the electrocardiogram by a minute dip (inverted P wave) preceding the upstroke of R. The P–R interval is now reduced to 0·027 second, and the interval between auricular and ventricular contraction is reduced almost to nothing; the two chambers are contracting simultaneously. The auriculo-ventricular interval (upper row of figures) and P–R interval (lower row of figures) are given in decimals of a second. Time marker=0·2 second (cf. Fig. 140B). (Lewis, *Mechanism and Graphic Registration of Heart Beat.*)

(1) If the S.A. node is destroyed or cooled, while the A.V. node is warmed, or if the right vagus is stimulated to depress the S.A. node and at the same time the left sympathetic is stimulated to heighten the excitability of the A.V. node, it is found that the *A.V. node becomes the dominating centre.* The impulse passes from this node at the same time to both auricles and ventricles, and the two chambers contract simultaneously. The P–R interval is greatly diminished; the P wave is inverted or is buried in the substance of the R wave (Fig. 165). In clinical records any reduction of the P–R interval below normal limits suggests that the impulse is arising in the region of the A.V. node. In view of the great structural resemblance between the S.A. and

A.V. nodes, it is not surprising to find that the A.V. node is possessed of such a high degree of rhythmicity.

(2) If the ventricles are suddenly cut off from the auricles, they may remain quiescent for a time as already explained (p. 284); they then resume beating, at first slowly and then more rapidly. If the ventricles are gradually dissociated from the auricles, the new rhythm attains its full rate very soon. The ventricular impulses now arise in the bundle of His just below the site of the injury so that the waves representing ventricular invasion (QRS) are of normal form (Fig. 162). This idioventricular rhythm is *not* controlled by the vagus.

(3) *Occasional* spontaneous beats may arise anywhere in the substance of the auricles or ventricles. True rhythmic function—the ability to initiate

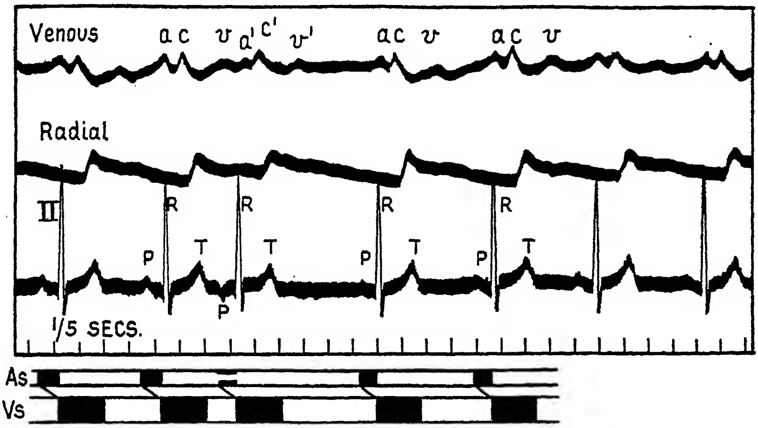


FIG. 166.—Auricular Extrasystole. Simultaneous Venous, Radial, and Electrocardiographic Curves illustrating Auricular Extrasystoles in a Patient, and a Diagram representing the Disturbed Mechanism (cf. Fig. 140B).

The third beat is the premature one; the P wave is inverted and follows very closely on the preceding T wave; the succeeding ventricular complex is normal. There is a premature pulsation at the wrist. The premature auricular wave *a'* precedes in a normal fashion *c'*, *v'*. A pause longer than the normal diastolic period follows the premature beat in this case. As, Vs=auricular and ventricular systole. (Lewis, *Mechanism and Graphic Registration of Heart Beat.*)

and *maintain* the heart beat over prolonged periods—is only present in the nodes and the junctional tissues.

Ectopic (Premature) Beats (Extra-Systoles.)—(1) AURICULAR.—If the *auricle* is stimulated during diastole after its refractory period has passed, it responds with a *premature* contraction. An impulse is transmitted to the ventricles, which contract too. The rate of recovery of the bundle is slower than that of the other parts of the heart, so that if an impulse reaches it prematurely, conduction along it is delayed; the P–R interval, therefore, is prolonged. The P wave shows a degree of abnormality which is greater the farther the stimulus is applied away from the S.A. node, because (cf. p. 245) the height and direction of the P wave are an indication of the way in which the excitation process has spread over the auricles. The P wave may thus be inverted, isoelectric, or abnormal in the details of its configuration. The ventricular complex, QRST, is normal. The next auricular impulse

arising in the S.A. node appears after a pause equal to the normal diastolic period or a little in excess of it (Fig. 166).

(2) VENTRICULAR.—If the *ventricle* is stimulated after the refractory period has passed, *i.e.* after the end of systole, it responds by contracting. As the ventricle obeys the " all-or-none " law, it responds to the maximum of its ability if the stimulus is adequate. The actual force of the contraction depends on the extent to which the ventricles have recovered from their previous contraction, and the degree of filling which has taken place. If the

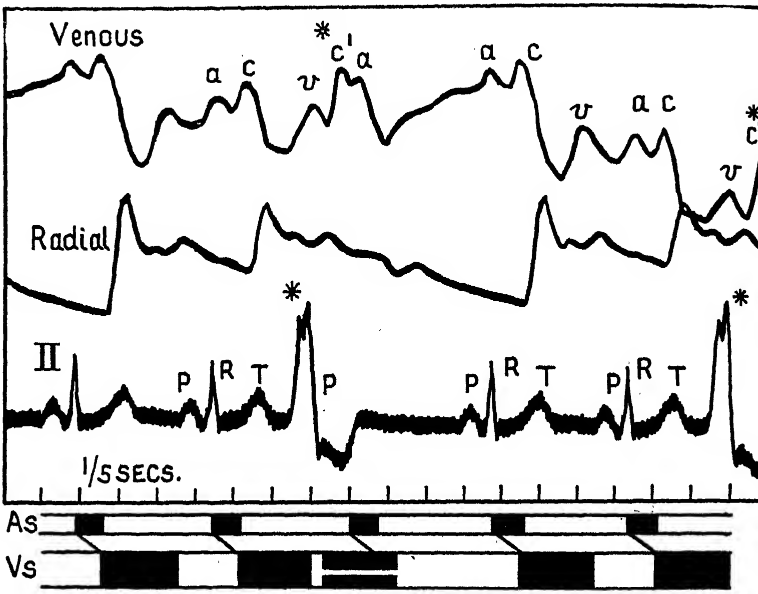


FIG. 167.—Ventricular Extrasystole. Simultaneous Venous, Radial, and Electrocardiographic Curves from a Patient, showing an Extrasystole arising in the Ventricle.

The diagram placed below the figure illustrates the mechanism of the heart during the period of the disturbance. As, Vs = auricular and ventricular systole.
The premature ventricular complex is abnormal in character; the P wave is buried in, instead of preceding, the ventricular complex. Similarly the *c′* wave is premature and precedes the *a* wave. The pause following the premature beat is longer than normal (compensatory pause); the auricular rhythm is undisturbed. (Lewis, *Mechanism and Graphic Registration of Heart Beat.*)

ventricle is stimulated early in diastole, the contraction it gives is feeble and may be insufficient to open the semilunar valves. The premature beat then only gives rise to the first heart sound and is not accompanied by arterial pulsation at the wrist. If the premature contraction occurs later in diastole, the ventricle may contract sufficiently forcibly to discharge its contents. As the ventricle has not had time to become completely filled, the output during this premature beat is less than normal, and the pulsation felt at the wrist is small (Fig. 167).

Following the extrasystole there is a long pause (*compensatory pause*). The duration of the " extrasystolic " cycle and the " returning " cycle (*i.e.* the cycle following on the premature contraction) is equal to two normal cycles. The reason for this compensatory pause is simple. The normal

auricular impulse which follows the extrasystole finds the ventricles in a refractory condition, and no response is obtained. The ventricles must therefore wait for the *succeeding* auricular impulse before contracting (Fig. 167).

Electrocardiographic Features.—The electrical record of the ectopic beat shows an abnormal ventricular complex which is not preceded by a P wave; the next P wave is generally "buried" within this ventricular complex (Fig. 167). The excitation process which arises in the new focus spreads radially over the surface of the ventricular muscle in all directions; it also penetrates the ventricular wall to reach the endocardium and so invade the specialized Purkinje tissue which transmits the excitation process rapidly

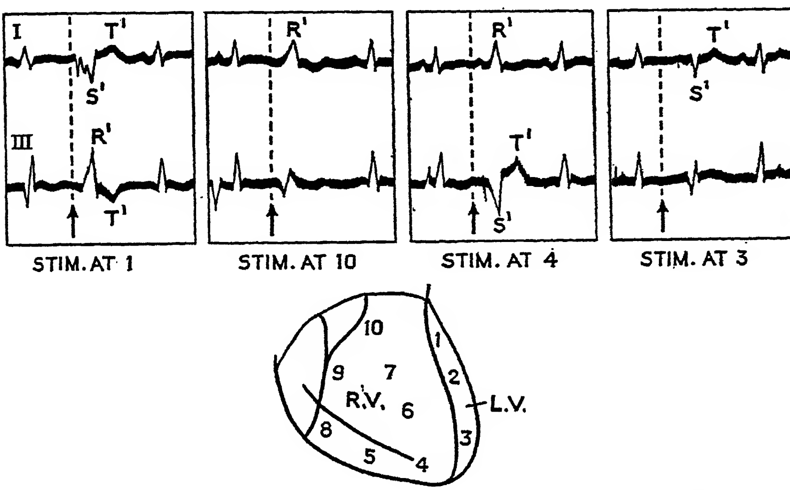


FIG. 168.—Electrocardiographic Records of Ectopic Ventricular Beats directly produced in Man.

The anterior view of heart shows the points which were stimulated on the surface of the ventricles; 1, 2, 3 are on the left ventricle, the others are on the right ventricle.
The electrocardiograms are from Leads I and III taken simultaneously, and show the curves resulting from stimulation of points 1, 10, 4, and 3. After Barker, MacLeod, and Alexander, *Amer. Heart J.*, 1930, 5.)

over its own side of the heart. The same changes occur later in the contralateral ventricle. As the time taken for the excitation process to affect the whole of both ventricles is prolonged the QRS will exceed 0·12 second in duration; as the pattern of invasion is abnormal the deflections of the QRS will be abnormal in appearance. Because of the delay in invasion the pattern of repolarization will be altered; the regions which were excited first may begin to repolarize before other regions have been fully invaded. Consequently the ST segment and T wave may be altered; there is no isoelectric portion of the ST segment; the T wave takes off from a level above or below the isoelectric line and usually has a direction opposite to that of the main wave of the QRS complex.

Though the electrical record enables an ectopic lead to be diagnosed as ventricular in origin it is difficult to determine from the record the specific region in the ventricle where the beat has arisen. Some guidance is obtained

from a study of the records produced by stimulating specific points on the human heart exposed at operation. The fact that the heart is exposed means that these records will not be identical with those produced by extrasystoles of similar site of origin with the chest closed. Some representative results are shown in Fig. 168. Four groups of reactions were obtained.

(i) From the anterior surface of the right ventricle (points 4, 5, 6, 7, 8, 9): the chief initial deflection is upwards in Lead I and downwards in Lead III.

(ii) From the conus region of the right ventricle (point 10): the chief initial deflections are upwards in all three leads.

(iii) From the left ventricle near the left auricular appendage (point 1): the chief initial deflection is downwards in Lead I and upwards in Lead III.

(iv) From all other points on the left ventricle (points 2, 3, and the posterior surface): the chief initial deflections are downwards in all three leads.

Right ventricular extrasystoles thus always gave an initial upward deflection (R′) in Lead I.; left ventricular extrasystoles always gave an initial downward deflection (S′) in Lead I.

Paroxysmal Tachycardia.—In this condition the heart may suddenly accelerate to 150 or 200 per minute. The new "pacemaker" may be situated in the auricle or in the ventricle (giving rise to auricular or ventricular tachycardia respectively). In one case studied, the output per minute fell from the normal level in that subject of 5·6 litres to 2·5 litres, and the output per beat from 77 c.c. to about 15 c.c.[1] This example illustrates how a very rapid pulse rate at rest may cripple the circulation. Diastole is too brief to allow proper filling of the heart, so that each beat only discharges a small amount of blood. The rest period is very short, the heart fatigues, and signs of heart failure ultimately appear.

Response of Auricles to Rapid Stimulation.[2]—As the frequency of stimulation of the auricles (dog) is increased the following sequence of events takes place:

(1) With rates up to 290 per minute the auricles respond regularly with contractions of uniform size; a *simple auricular tachycardia* is thus set up.

(2) When the auricles are stimulated at 290–380 per minute, a condition called the *partially refractory state* develops. By this is meant that at any given moment some auricular muscle fibres are responsive, while adjacent fibres are refractory. In response to each stimulus at this rate some groups of fibres react and others fail to respond. If a large part of the auricles responds, a big beat is obtained; when a small area reacts, a small beat results. The fibres which respond at one beat are not sufficiently recovered and fail to react to the succeeding stimulus, so that alternate large and small contractions result.

As the refractory period of the ventricles is longer than that of the auricles, they develop a 2 : 1 response when stimulation occurs at 350 per minute. The A.V. node cannot transmit more than 270 impulses per minute.

(3) With still more rapid rates of stimulation, *e.g.* 380–450 per minute, or over, the length of the refractory period of the auricular muscle is *increased*, and is *longer than the interval between succeeding stimuli*. A state of 2:1 response

[1] Barcroft, *Heart*, 1922, *9*, 419.

[2] For discussion of this condition, auricular flutter and auricular fibrillation, see Lewis, *Brit. med. J.*, 1921, *1*, 551, 590. Lewis and associates, *Heart*, 1920, *7*, 191, 247, 293; 1921, *8*, 141, 193, 311; 1922, *9*, 55, 207.

of the auricles then results, *i.e.* every other stimulus finds the auricles in a completely refractory state, and so a contraction results from alternate stimuli only.

Efficient *conduction* in the auricles depends on the *excitation process meeting responsive tissue which can transmit the impulse farther.* When the auricles are stimulated at high rates, *conduction in the auricles is slowed*; considerable areas of the auricular wall may be in a refractory state when the excitation process reaches them. These form areas of obstruction, and the impulse instead

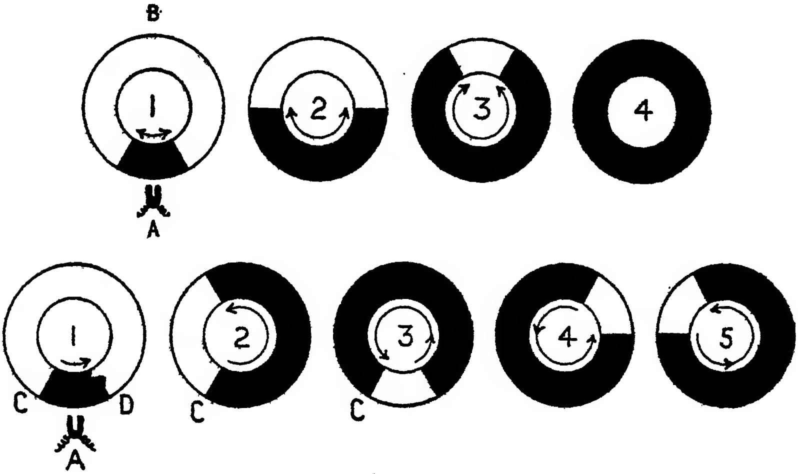


FIG. 169.

UPPER LINE (1–4).

Diagram illustrating Progress of Single Wave passing through a Ring of Muscle as a result of stimulating it at point marked by Electrodes (A). Black portion of ring represents refractory state.

LOWER LINE (1–5).

Diagram to illustrate the Establishment of a Circus Movement in a Ring of Muscle. The ring is stimulated at A and the wave spreads to C and D. At C it is blocked, but from D it continues round the ring. When it arrives at C, the refractory state has passed off and so it continues to travel round the circle. (Lewis.)

of flowing smoothly must wind about irregularly to seek out portions of tissue which have recovered. The rate of conduction in the auricles when they are stimulated at high rates may fall to 500 mm. per second (normal=1000 mm.).

Following brief stimulation at these very high rates, auricular flutter or auricular fibrillation may set in (p. 292).

Circus Movement.—If a ring of muscle is taken, and a point upon it, A (Fig. 169, upper line), is stimulated, a wave of excitation develops which spreads equally in both directions till it reaches the opposite side of the ring at B. While the muscle is in the grip of the excitation process, it is completely refractory, so when the crests of the excitation waves reach the point B, they are faced by unresponsive tissue, and can proceed no farther. The wave of excitation is thus arrested, and the active state gradually passes off from the whole ring of muscle.

If the rate of stimulation is increased, we have already learnt (p. 290) that at a certain critical rate, just before 2 : 1 responses are obtained, a condition called the partially refractory state develops. Certain parts of the muscle have fully regained their responsiveness, others are still refractory when each stimulus falls on the muscle. In the case of a ring of muscle (Fig. 169, lower line), let the point A be stimulated so that it responds and an impulse spreads out from it. At C, let it be supposed that it finds itself faced by an area which has not recovered sufficiently; the impulse in this locality is consequently extinguished. The region D, however, happens to be responsive; the impulse will therefore spread in that one direction only. If conditions happen to be suitable, it finds responsive tissue before it, and travels on as a single-crested wave right round the ring. When the point C is reached (which was refractory before) it is found to be responsive, and so the wave travels on to the original starting-point A. This point too has recovered from its refractory condition, and so a fresh cycle begins. The wave of excitation can thus go on circulating round the ring for minutes or hours. The term *circus movement* is used to describe the wave which travels continuously round and round in one direction (Fig. 169, lower line).

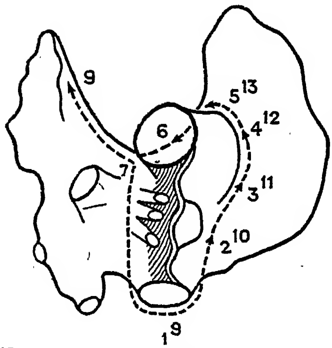


FIG. 170.—Experimental Auricular Flutter. Outline of dog's auricles, showing the supposed paths of the excitation wave (broken lines) during auricular flutter. The numbers indicate the order in which the various parts of the musculature were thought to receive the excitation wave. (Lewis.)

To establish a circus movement the following conditions must be fulfilled:

(i) An area of refractoriness must be present to prevent the passage of the impulse in one direction.

(ii) The originally refractory area must have recovered when the excitation process reaches it for the second time.

According to Lewis a circus movement in the auricles is responsible for auricular fibrillation and flutter; this view has, however, been refuted by recent work.

Auricular Flutter.—In this clinical condition the auricles beat at 230–350 times per minute with considerable regularity and for prolonged periods. The bundle of His cannot transmit impulses at such high frequencies; the ventricular rate is generally half or even a quarter of the auricular rate and shows a regular or an irregular rhythm.

Auricular flutter can be produced experimentally in the following manner. The auricles are stimulated by induction shocks at 300–600 times a minute. For a few seconds the auricles respond very rapidly. The stimuli are then withdrawn. Instead of resuming their original rate, the auricles may continue to beat rapidly, but not at the rate at which the induction shocks had been thrown in. According to Lewis a circus movement has been set up which travels continuously round a ring of auricular tissue, *e.g.* the tissue joining the openings of the inferior and superior venæ cavæ (Fig. 170). Instead of the auricles being stimulated, as normally, from a single point, the sino-auricular node, there is a sort of circulating centre of impulse formation. As the wave

travels round the ring, the excitation process spreads out from the ring to activate the outlying parts of the auricular muscle. The auricles contract once, each time the wave completes one circuit of the ring.

Experimental re-examination of the question has shown, however, that a circus movement is *not* set up. Auricular flutter is simply a high grade auricular tachycardia, the impulses being set up at high frequency (usually 230–350 per minute) from a single ectopic focus.

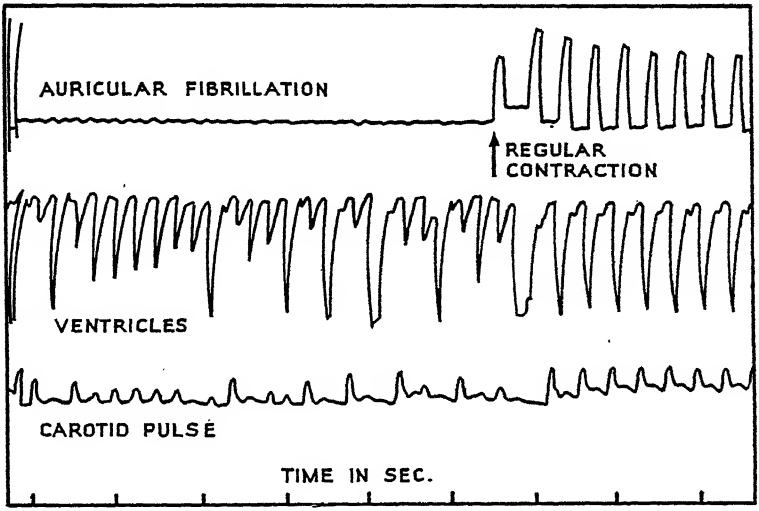


Fig. 171.—Experimental Auricular Fibrillation.

Experiment on a Dog; chest open.
Records from above downwards: contraction of auricles, contraction of ventricles, carotid pulse, time in seconds.
During the first part of the record the auricles are in a state of fibrillation (fine irregular oscillation in diastolic position); the beat of the ventricles is completely irregular in force and rate; the carotid pulse is also completely irregular.
At the arrow, regular auricular contractions spontaneously returned; the ventricles respond regularly to the auricular beats, and the carotid pulse likewise becomes regular. (Redrawn from Lewis, *Mechanism and Graphic Registration of Heart Beat*, 1925.)

Auricular Fibrillation.—This condition can be induced experimentally by transient electrical stimulation of the auricles (at 300–600 times per minute) or by local application of aconitine. High speed colour films of the auricles in the dog have been taken at 2000 frames per second and later projected at 8 frames per second; the events of one second in the auricle unfold themselves in 4 minutes on the screen. A magnification of 100-fold is simultaneously obtained.[1] These records show that in auricular fibrillation, two types of muscular activity occur:

(i) Moderately irregular contractions, visible to the naked eye, repeated at a rate of 400–600 per minute and not uniform in strength. These are shown in Fig. 171; they are associated with small irregular, electrical oscillations labelled *f*, *f* (Fig. 172), which replace the P wave. There is *no* evidence of a circus movement.

[1] Prinzmetal *et al.*, *Circulation*, 1950, *1*, 241; *J. Amer. med. Assoc.*, 1951, *146*, 1275.

(ii) Much smaller, very irregular, contractions occuring at very high speeds. Frequencies of 800–1600 per minute are recorded by direct leads; more sensitive methods are claimed to record oscillations at frequencies up to 40,000 per minute. These excursions are rightly called minute or "M"; by contrast, and only by contrast, the tiny excursions referred to in (i) are called "L" (large).

As in flutter, so in fibrillation, the A.V. bundle does not transmit all the impulses reaching it. The ventricular rate is much less than that of the auricles (*e.g.* only 130 per minute), and the beats are completely irregular and bear no constant relationship to one another. The beats may similarly be large or small, depending on how much blood is present in the left ventricle at the beginning of each systole. This rapid irregularity greatly interferes with the output per minute, which falls to less than 4 litres, and each beat of the heart discharges only 30–40 c.c. Auricular fibrillation commonly

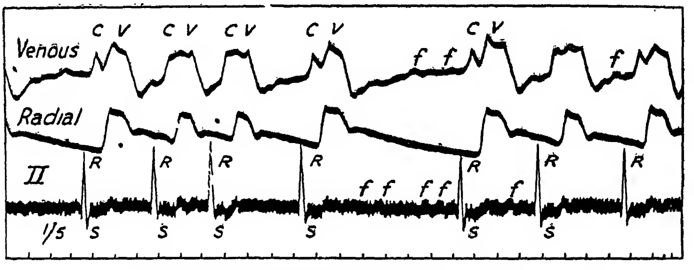


FIG. 172.—Auricular Fibrillation in Man.

Simultaneous Records of Jugular Venous Tracing, Radial Curve and Electrocardiogram (Lead II). The P waves are replaced by small irregular oscillations (f, f). The ventricular complexes are of supraventricular origin and occur at completely irregular intervals (note intervals between R waves).
The radial curve displays the gross pulse irregularity. No *a* waves are present in the jugular curve. (Lewis, *Mechanism and Graphic Registration of Heart Beat.*)

develops in subjects with mitral stenosis, in severe hyperthyroidism and in degenerative states of the myocardium.

Typical graphic records from a case of clinical auricular fibrillation are shown in Fig. 172.

Ventricular Fibrillation may be produced as follows:

(i) Application of a strong constant or faradic current to the ventricles.

(ii) Ligation of a coronary artery (cf. p. 238).

(iii) Administration of toxic substances, including excessive doses of digitalis.

(iv) When low percentages of *chloroform* vapour are inhaled, the irritability of the ventricles is enhanced to a precarious degree. Injections of minute doses of *adrenaline*, stimulation of a sensory nerve, or some inconspicuous and unnoticed influence, or even a fresh dose of stronger vapour, may precipitate fibrillation.[1]

(v) Not infrequently following simultaneous division of both sinus and aortic nerves.

SEQUENCE OF EVENTS.—(i) Ventricular extrasystoles arise in one or many foci. (ii) Rapid, almost regular, small undulations of the ventricles occur.

[1] Levy, *Heart*, 1913, 4, 319.

(iii) Fibrillation sets in, and co-ordinated contractions of the heart fibres cease. The heart dilates and the blood pressure falls to zero. On inspection the ventricles display minute quivering movements; one area is contracted while an adjacent one is relaxed. The electrocardiogram shows large irregular oscillations.

Clinically, fibrillation of the ventricles has been recorded in the *dying* heart. The earlier stages (i) and (ii) have been recovered from. Ventricular fibrillation is the mechanism of sudden and unexpected death, preceded by loss of pulse, pallor passing to cyanosis, and a few gasping respirations.

HEART FAILURE. CIRCULATION IN VALVULAR DISEASE[1]

Heart Failure.—The efficiency of the heart in maintaining an adequate circulation is best judged from the patient's symptoms, the results of routine clinical examination, and most important of all, by the response of the patient to various degrees of exertion. An inadequate blood supply to the heart muscle manifests itself in *pain* on exertion (p. 750). Failure of the *left ventricle* leads to a fall of arterial blood pressure and decreased blood flow to the principal organs, especially the brain and kidney, with the expected results (p. 460) ;if the right ventricle continues to beat normally, engorgement and œdema of the lungs results (p. 116). Failure of the *left auricle* leads to its enlargement and to obstruction to the outflow from the lungs, again with resulting pulmonary congestion and œdema. These lung changes lead to a decrease in the *vital capacity* (p. 363).

Failure of the *right ventricle* leads to engorgement of the *right auricle* with resulting rise of general *venous pressure* (p. 319), swelling of the liver, *œdema* (p. 111), and serous effusions into the peritoneum or pleura. *Cyanosis* and *dyspnœa* are very common manifestations of heart failure, which are fully discussed on pp. 451 and 460. (These references should be carefully consulted.)

Determinations of the *circulation time* (p. 262) confirm but generally add little to the results of clinical examination. Determinations of cardiac output in heart failure have yielded little information of use. It must be remembered that the cardiac output methods available in man really measure the *pulmonary* blood flow, which equals the right (and normally also the left) ventricular output. In left-sided valvular disease (*e.g.* aortic incompetence, mitral incompetence) the cardiac output as determined by the Fick method (p. 278) is a measure of the right ventricular output and therefore of the venous return and of the actual blood flow to the tissues (*i.e.* the " peripheral blood flow ") ; it thus measures the *net* output of the left ventricle into the arteries and not the *actual* output. The *size* of the heart as determined by X-ray examination is a better index of the stroke volume in these valvular disorders. It is disappointing moreover to find little relationship between the peripheral blood flow (*i.e.* the cardiac output as measured by the usual techniques) and the state of the circulation as judged clinically. The cardiac output may be apparently unreduced when clinically circulatory failure is manifest; no explanation is available for this strange discrepancy.

[1] Lewis, *Diseases of Heart*, 4th edn., London, 1946. Fishberg, *Heart Failure*, 2nd edn., 1940.

The significance of electrocardiographic records is considered on pp. 251 *et seq.* and 282 *et seq.*

Experimental Mitral Incompetence.—In experimental mitral lesions, little if any regurgitation into the auricle occurs during the isometric contraction phase of the ventricular systole. This is due to the shortness of this phase (0·05 second), and also to the fact that left intraventricular pressure during this phase of the cycle is relatively low, ranging up from a few mm. Hg to only 60 mm. Hg (Fig. 173).

The principal backflow occurs during the *ejection* phase (when left ventricular pressure is very high), and it continues for 0·08 second into diastole. The intraventricular pressure then falls below the intra-auricular, and the auriculo-ventricular (A.V.) valves open. The auricular volume and pressure increase markedly throughout the period of regurgitation. Because of the regurgitation, the systolic discharge of the left ventricle into the aorta is at

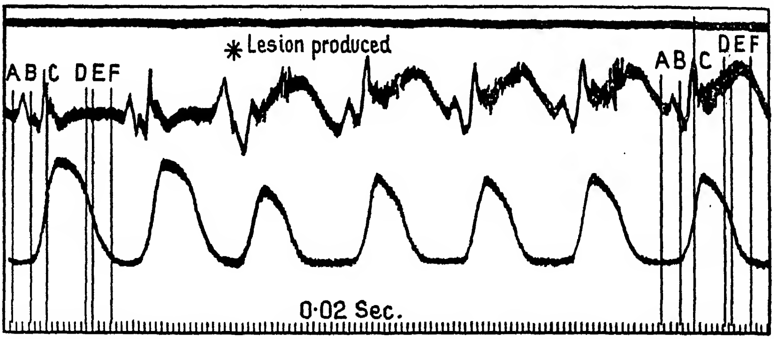


FIG. 173.—Simultaneous Records of Auricular (upper) and Ventricular (lower) Pressure Changes in Experimental Mitral Incompetence.

AB=Auricular systole; BC=Second positive wave in auricle and isometric ventricular contraction CD=Ejection phase of ventricle; E=Closure of aortic valves; EF=Isometric ventricular relaxation. At * mitral incompetence produced. Note that AB, BC, in auricular record are little altered. *During ventricular ejection,* CD, *auricular pressure is markedly raised by regurgitating blood.* Regurgitation continues through early part of diastole. Time in 0·02 second. (Wiggers and Feil, *Heart.*)

first reduced. The initial result, therefore, of an experimental mitral regurgitation is that the systolic and diastolic blood pressure fall and the pulse pressure is decreased. Compensation, however, is soon effected.

When the A.V. valves open, the left auricle, which is engorged with blood (which has come from the right heart and also through the mitral leak from the left ventricle), discharges an *increased* volume into the left ventricle. This results in greater initial length of the ventricular fibres and more forcible contraction. Some of the ventricular blood regurgitates into the left auricle, but enough is pumped into the systemic arteries to maintain an adequate circulation.

The augmented volume of blood in the left auricle is accommodated by expansion of the left auricle and pulmonary veins. When the heart is well compensated the pressure in the pulmonary bed only rises temporarily during the ventricular ejection phase when regurgitation is occurring.

Disease of Mitral Valves.—Clinically, mitral valvular disease almost

invariably leads to the occurrence of both *incompetence* and *stenosis* of the valve orifice. *Stenosis* interferes with the outflow of the blood from the left auricle, both during the passive outflow which occurs early in diastole and the active outflow resulting from auricular systole (p. 258). Incompetence leads to *regurgitation* of blood from the left ventricle into the left auricle during ventricular systole as described above. The left auricle is therefore overdistended during diastole when it receives its normal inflow from the pulmonary circuit. The greater stretch of the muscle fibres leads as usual to greater force of contraction; a larger volume of blood (both that which has regurgitated and that which has newly arrived) is discharged into the ventricle. When compensation is perfect enough extra blood is sent into the left ventricle to make up for the volume which regurgitates; the actual ventricular outflow into the arterial system is at this stage undiminished although at the expense of greater ventricular effort. If the stenosis becomes more severe, ventricular filling may be diminished and the output into the aorta is reduced. If regurgitation becomes very free, a large fraction of the ventricular inflow leaks back into the auricle and again the flow into the arterial system suffers. As compensation fails there is increasing dilatation of the left auricle with the consequences detailed on p. 460. If auricular fibrillation sets in the co-ordinated contraction of the auricles ceases and the ventricles can only fill *passively* in diastole when the auricular pressure exceeds that in the ventricles. The rapid irregular contractions of the ventricles interfere further with the cardiac output which is seriously diminished. [For electrocardiographic changes, see Fig. 147, A, and p. 247; heart sounds, p. 267 and footnote.]

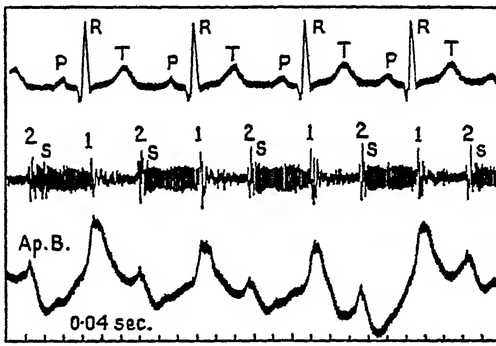


FIG. 174.—Heart Sounds in Aortic Regurgitation. (Braun-Menendez and Orias, *Heart Sounds in Health and Disease*, London, 1939.)

Records from above downwards are: electrocardiogram; heart sounds (1=1st sound, 2=2nd sound, S=aortic diastolic murmur); apex beat pulsations (Ap.B.).

Aortic Regurgitation.—In this condition some of the left ventricular output leaks back during diastole from the aorta into the ventricle; a characteristic diastolic murmur is set up (Fig. 174). The left ventricle is simultaneously being normally filled from the left auricle; at the end of diastole the ventricle is excessively distended, the muscle fibres are more stretched and accordingly the force of contraction is increased; the output per beat correspondingly rises. The chronic overwork of the left ventricle gradually leads to hypertrophy of its fibres (p. 277). Owing to the raised output the systolic blood pressure is high (up to 180 mm. Hg). If compensation is perfect the *net* ventricular output, *i.e.* the blood supply to the tissues ("peripheral blood flow") is undiminished; to accomplish this end the output of the left ventricle per beat has to be increased by a volume

of blood equal to that which regurgitates during each diastole back again into the ventricle (so-called *shunt blood*), and which, of course, never passes to the periphery. The diastolic pressure is low (and the pulse pressure therefore raised); this is mainly due to the leak back into the ventricle. It is also suggested that the large cardiac output may set up an exaggerated sino-aortic reflex at the height of systole with resulting marked transient arteriolar dilatation. This reflex leads to a rapid escape of blood from the arterioles into the capillaries; it also contributes to the collapse of the pulse (*infra*).

Cases of *arterio-venous aneurysm, i.e.* where there is a large abnormal communication between an artery and a vein, present many of the arterial phenomena of aortic regurgitation. In both conditions blood leaks out of the arteries by other than the natural pores provided by the capillaries. Certain differences are naturally to be expected, because in aortic regurgitation the leak is confined to diastole and in the aneurysm the leak occurs throughout the cardiac cycle. The following arterial phenomena are *common to both conditions* and are abolished by closing the anastomosis in arterio-venous aneurysm; they are therefore *due to the abnormal leak* which is present:

(i) Low diastolic pressure (*supra*).

(ii) "*Water hammer*" character of the pulse, *i.e.* a sudden impact is felt by the palpating hand. This sensation is due to the rapid upstroke in the initial phase of the primary wave, which may also be of greater amplitude than normal. The phenomenon is much exaggerated by holding up the arm vertically.

	Time from Beginning of Upstroke to Half-way Point.	
Horizontal position of arm .	Normal, 0·039 sec.	Aort. Reg., 0·027 sec.
Vertical position of arm .	Normal, 0·045 sec.	Aort. Reg., 0·017 sec.

The speedy ascent is due to the abnormally low diastolic pressure. As ventricular ejection meets with very little resistance, the blood leaves the ventricle very rapidly, and the aortic pressure rises with abnormal speed. *Visible pulsation* in the arteries of the neck is due to the same reason, *i.e.* sudden distension of the underfilled vessels.

(iii) *Collapse of the Pulse.*—After the sudden impact just described, the pulse rapidly falls away. Part of the fall occurs during *systole* (when *no* regurgitation is occurring), *i.e.* it precedes the dicrotic notch; this is attributed to a reflex vasodilatation of the peripheral arterioles (cf. p. 261). There is also a steep fall which *follows* the dicrotic notch and therefore occurs during diastole; this is due to the reflux of blood which is then occurring into the ventricle.

A case showing all the arterial signs enumerated above is described by Lewis as one of "free" aortic regurgitation. The degree of leak must be equal to that present in arterio-venous aneurysm when the same arterial phenomena are present. The leak, it is found, may equal or exceed 50% of the total cardiac output.

Blood Pressure in Leg and Arm.—In aortic regurgitation there is *hypertonus* of the walls of the main leg arteries, which diminishes the lumen of these vessels and so keeps the blood out of the dependent parts. When the blood pressure is determined with the subject in the horizontal position, the

apparent systolic pressure is higher in the leg than in the arm as the spasm of the wall of the leg vessels must be overcome before obliteration of the vessel can begin. If the limb is warmed the hypertonus disappears, and the systolic pressure readings in leg and arm become the same.

Capillary pulsation is present in aortic regurgitation, and is seen as a rhythmic expansion of the capillary loops with each beat of the heart; the pulsation affects the minute skin venules, too. Capillary pulsation is a *physiological* phenomenon and occurs in the skin and mucous membranes whenever the arterioles are sufficiently dilated, provided that the pulse pressure is normal and capillary and venule tone is not increased. It is to be commonly observed in young people at room temperatures of 17°–20° C., after soaking the hand in water at 45° C., and in the facial flush produced by amyl nitrite. Capillary pulsation should therefore be regarded as associated with dilated arterioles, which permit the pulse to reach the minute skin vessels, rather than with any particular morbid condition.

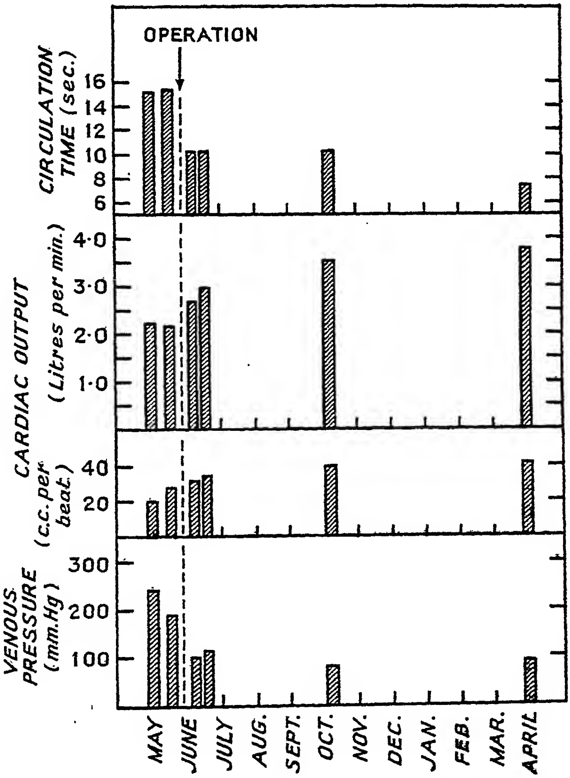


FIG. 175.—Circulatory Changes in Constrictive Pericarditis and the Effects of Pericardial Resection. (Stewart and Heuer, *Arch. int. Med.*, 1939, *63*, 507.)

Circulation time in seconds: a decrease in the time indicates increased cardiac output.
Cardiac output: in litres per minute and c.c. per beat.
Venous pressure in cm. H_2O: a fall of pressure indicates decreased venous congestion.
Initially there was marked enlargement of the liver.

It should be noted that the regurgitant stream in aortic incompetence raises the intraventricular tension in diastole and thus impedes the filling of the left ventricle through the mitral orifice. The filling of the *coronary arteries* is impaired because it normally occurs mainly during diastole and depends on the blood pressure during diastole which in aortic regurgitation is lowered. The nutrition of the heart muscle suffers, and dilatation results.

The changes that take place in the circulation when failure of compensation occurs [1] in aortic regurgitation are discussed on p. 460.

Constrictive Pericarditis.—This clinical condition (often called *adherent pericardium*) illustrates well the results of external mechanical interference with the heart's action. The pericardium is thickened and firmly adherent to the surface of the heart and to the mediastinal structures. This indistensible casing interferes with diastolic filling of the heart; it likewise hampers the shortening of the muscle fibres during systole. The cardiac output is consequently reduced, the circulation time is prolonged, and the venous pressure is raised (Fig. 175). Resection of the thickened pericardium leads to great circulatory amelioration.

REGULATION OF BLOOD PRESSURE.[2] PERIPHERAL CIRCULATION

The blood pressure is the lateral pressure exerted on the walls of the vessels by the contained blood. The *systolic* arterial pressure is the maximum pressure in the arteries during systole; the *diastolic* pressure is the lowest pressure during diastole. The *pulse pressure* is the difference between the diastolic and systolic pressure.

Measurement of Arterial Blood Pressure in Man.—The *sphygmomanometer* is used; the rubber bag (enclosed in an unyielding cuff) is fixed round the upper arm and connected with the pressure pump and a mercury or spring manometer. The pressure in the bag is rapidly raised to 200 mm. Hg to obliterate the brachial artery and to overcome any spasm which may be present; the subsequent procedure depends on whether the palpatory or auscultatory method is being used.

(i) Palpatory Method.—The cuff pressure is lowered until the pulse can just be detected at the wrist. This is the measure of the systolic pressure; diastolic pressure cannot be determined by this method.

(ii) Auscultatory Method.—A stethoscope is placed over the brachial artery below the cuff. The cuff pressure is lowered and a series of characteristic sounds is heard.

First phase: suddenly a clear sound is heard. The systolic blood pressure is the pressure at which the sound is *first heard*. The sound continues to be heard while the pressure is lowered by about another 15 mm.

Second phase: the sound becomes murmurish in character; it can be heard while the pressure is lowered by about another 15 mm.

Third phase: the murmur is replaced by a sound which becomes progressively louder and more "banging" in character during the next 20 mm. of fall of pressure.

Fourth phase: the sound suddenly becomes soft and muffled. The point at which the sound *begins to fade* is the accepted criterion of the *diastolic pressure*.

[1] Altschule, *Medicine*, 1938, *17*, 75.

[2] McDowall, *Control of Circulation of Blood*, 1938. Assoc. Res. nerv. ment. Dis., *Circulation of Brain and Spinal Cord*, 1938. Franklin, *Brit. med. J.*, 1951, i, 1343, 1410.

Fifth phase: cessation of all sound. In many cases, however, total cessation of the sound does not occur.

For clinical purposes, in addition to "casual" readings, it is important that *basal* readings should also be taken: the patient is at rest in the recumbent position or in the sitting position with the artery at heart level; *readings are repeatedly taken until a steady minimal level is obtained.*

Normal Blood Pressure.—During childhood the systolic pressure gradually increases. It ranges from 75 to 90 mm. Hg during infancy, from 90 to 110 mm. in childhood, and 100 to 120 mm. about puberty. The diastolic pressure ranges round 50 mm. during the first few years of life, and after that until puberty it remains fairly constant at 60 mm.

In adults the *average* normal systolic pressure is usually stated to be about 125–130 mm. Hg with a range of $\pm$ 15, *i.e.* values between 110 and 145 mm. are within the range of the normal. In older people it is between 140 and 150 mm. (Fig. 176).

A study of over 10,000 apparently normal subjects of all ages, in many cases for periods of 10 years, has led to slightly different conclusions.[1] The usual normal range of systolic pressure was 90–120 mm. Hg; that of diastolic pressure was 60–80 mm. The "mature" pressure was reached in adolescence and *did not rise with age*. The upper limits of normal systolic and diastolic pressure were placed at 140 and 90 mm. respectively. Most clinicians regard a basal systolic blood pressure which regularly exceeds 150 mm. Hg as indicating a deviation from normal.

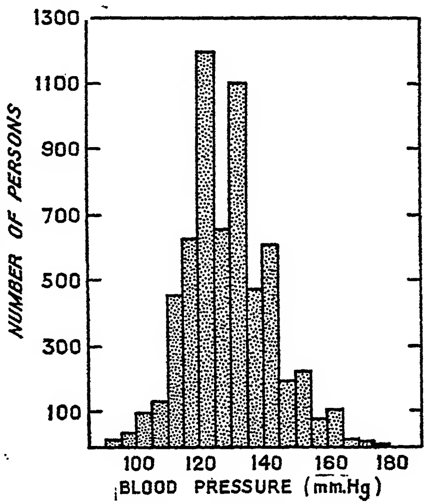


Fig. 176.—Distribution of Blood Pressure (Systolic) in Normal People. (After Burn, *Proc. roy. Soc. Med.*, 1934, *28*, 15).

The *systolic pressure* may fluctuate with physiological conditions such as exertion (p. 435), mental state, sleep, or meals.

The *diastolic pressure* is a measure of the peripheral resistance, and depends mainly on the tone of the arterioles. It is less subject to temporary fluctuations than is the systolic pressure. It is normally 35–50 mm. Hg lower than the systolic pressure.

The *pulse pressure* is the rise of pressure produced by ventricular systole. Though there is a general correspondence between the size of the pulse pressure and the output of the heart in the sense that both increase or decrease synchronously, there are no regular quantitative relations between the two. The same pulse pressure may at different times and in different individuals correspond to systolic discharges of very different volumes.

Graphic Recording.—In animals the blood pressure can be recorded

[1] Robinson and Brucer, *Arch. int. Med.*, 1939, *64*, 409.

graphically by inserting a cannula into the carotid artery and connecting it with a mercury manometer provided with a float and writing-point which records on a moving smoked surface. The cannula and tubing are filled with an anticoagulant fluid. These records show an oscillation with each heart beat. The real pulse-pressure difference is not accurately recorded, as owing to its inertia the mercury does not rise to full systolic level or fall to the real diastolic level. With a *slowly* beating heart there is more time to overcome the inertia of the mercury, and the oscillations are larger; with a rapidly beating heart the oscillations are smaller in extent. The size of the oscillation is *not* a reliable indication of the cardiac output or force. Commonly, too, a variation with the phases of *respiration* can be seen (Fig. 186 and

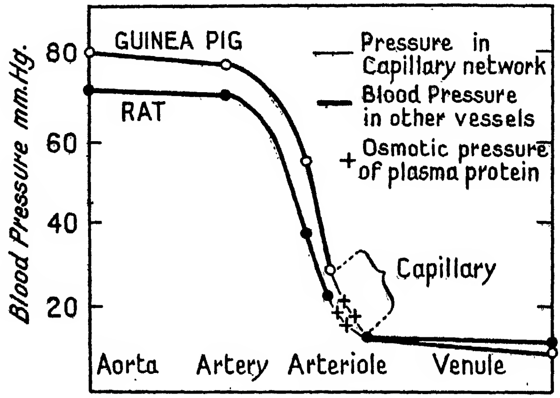


FIG. 177.—Pressure Gradient in Different Parts of the Vascular Bed.

Note that the greatest fall of blood pressure occurs in the arterioles. The osmotic pressure of the plasma proteins is indicated by the +, and is intermediate in value between the pressures in the arteriolar and venous capillary. (After Landis, *Amer. J. Physiol.*)

p. 313). Direct blood pressure recording from the interior of arteries has also been carried out in man (Fig. 186).

Factors controlling (Arterial) Blood Pressure.—The blood pressure varies as the *product of the cardiac output and the peripheral resistance.* To a minor degree it also depends on the *elasticity* of the large arteries (p. 262). The cardiac output depends on the venous return, the force and the rate of heart (for full discussion see pp. 274 *et seq.*). Other factors remaining unaltered, an increase in cardiac output raises, and a fall in output lowers, the blood pressure. The peripheral resistance will now be considered.

Peripheral Resistance.—The peripheral resistance is found chiefly in the tonically contracted *arterioles* (especially those of the splanchnic area and the skin), and to a less extent in the *capillaries.*

The pressure gradients in the various parts of the vascular bed have been determined directly in several species (Fig. 177). Normally, the greatest fall of pressure takes place while the blood is flowing through the ***arterioles***; but a considerable further fall occurs in the capillaries too.

The frictional resistance encountered by fluid passing through narrow tubes depends on (i) its *viscosity*; (ii) size of the *lumen* of the tube; (iii) *velocity* of flow.

The factors determining the viscosity of blood are discussed on p. 135. A decrease in viscosity lowers the resistance, an increase in viscosity raises the resistance. If the viscosity of water is taken as 1, plasma is 1·55; blood samples with corpuscular hæmatocrit values of 25, 35, 45, 55 and 65% have viscosities of 1·7, 1·9, 2·1, 2·6 and 3·0 respectively. The greatly increased viscosity of the blood in erythræmia (p. 168) accounts in large measure for the associated hypertension.

As the capillaries are so much narrower than the arterioles, it may be wondered why the resistance is principally met with in the latter vessels; the reason is that rapidly flowing fluid meets with much more resistance than slowly flowing fluid, and the velocity in the arterioles is many times higher than in the capillaries.

The part played by the arterioles is so much better known than that of the capillaries that the latter are often left out from discussions of the peripheral resistance, but doubtless they participate in many reactions. The significance of capillary tone is discussed on p. 320. In some conditions arteriolar dilatation is associated with capillary contraction: the major part of the peripheral resistance may then be located in the capillaries.

Fig. 178 summarizes the factors which regulate the tone of the arterioles.

The Vasomotor Centre.—The degree of contraction (*tone*) of the arterioles (and probably also of the capillaries and veins) is under the continuous control of the vasomotor centre which is situated in the floor of the fourth ventricle at the level of the apex of the calamus scriptorius. The centre gives rise to fibres which pass down the cervical cord (in the white matter of the lateral columns (Fig. 687) to all the thoracic and the upper two lumbar segments to end in the connector cells of the sympathetic which lie in the lateral horn of grey matter. The vasoconstrictor fibres then pass out in the sympathetic nervous system (described in detail at pp. 707 *et seq.*) to the various parts of the body.

The vasomotor centre is continually sending out excitor impulses to the arterioles to maintain their normal degree of tone. The rate of impulse conduction in vasomotor nerves is slow as the preganglionic are B fibres, and the postganglionic are C fibres (p. 492). The frequency of the impulses in these fibres varies with the degree of activity of the vasomotor centre. Thus asphyxia increases the frequency of the discharge; depression of the vasomotor centre may cause a temporary arrest of the discharge in sympathetic nerves. There is a variation in the sympathetic discharge with the phases of respiration; the maximum discharge occurs towards the end of inspiration owing to an *irradiation from the respiratory to the vasomotor centre* which is situated in very close relationship to it (cf. p. 271 and Fig. 156).

When the medulla oblongata is trans-sected at the level of the calamus scriptorius, a fall of blood pressure occurs down to about 40 mm. Hg, because the tonic impulses from the vasomotor centre are cut off. Section of the spinal cord above the level of the first thoracic segment produces an equally severe depression of the blood pressure, because the vasomotor centre has been cut off from all the sympathetic connector cells in the spinal cord. Section below the level of the second lumbar segment produces a minimal effect,

because all the vasomotor fibres have already passed out of the cord (cf. p. 690).

Following division of the spinal cord, the blood pressure finally recovers owing to the connector cells in the cord functioning independently as subsidiary vasomotor centres (p. 691).

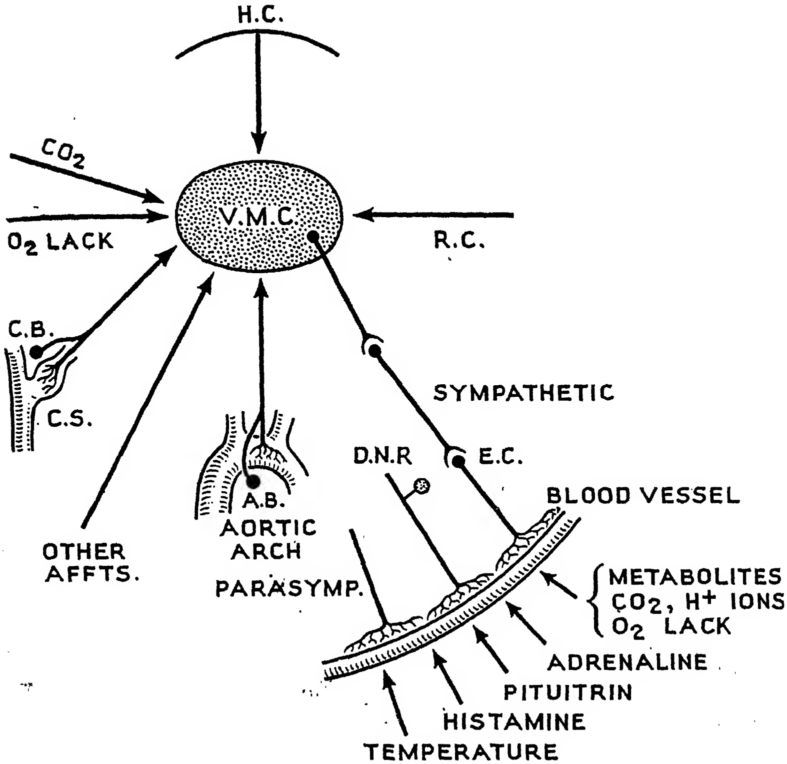


Fig. 178.—Regulation of Arterial Tone.

V.M.C.=Vasomotor centre; R.C.=Irradiation from respiratory centre; C.S., C.B.=Carotid sinus and carotid body; C.C., E.C.=Connector and excitor cells of sympathetic nervous system; D.N.R.= Dorsal nerve root dilator fibres; Parasymp.=Cranial or sacral autonomic dilator fibres; A.B=Aortic. body; H.C.=Higher centres (cerebral cortex and hypothalamus).

The activity of the vasomotor centre cannot be determined merely by recording the general blood pressure. A rise in blood pressure may be due either to increased peripheral resistance or increased cardiac output (or both); if the peripheral resistance is halved and the cardiac output doubled, the blood pressure remains unchanged. The state of the blood vessels and the peripheral blood flow in various regions can be studied by the methods described below; particular attention will be paid to methods which are used in man.

Methods and Results of Studying Peripheral Circulation.—(1) Plethysmography.—The organ (in animals, limb, kidney, spleen,

intestine, etc., or in *man*, part of a limb such as hand, forearm, or hand plus forearm) is enclosed in an airtight vessel (*plethysmograph*, *oncometer*) which is connected with a recording apparatus (Fig. 179). Changes in the volume of the organ are due to alterations in the amount of contained blood, which in turn depends on the calibre of the blood vessels. Expansion of the organ represents vasodilatation (Fig. 187, A), shrinkage of the organ indicates vasoconstriction. These changes may be *actively* produced by an alteration of local vascular tone or *passively* produced by a rise or fall of arterial or venous blood pressure. The plethysmograph record does not give information about the particular section of the vascular system involved (*i.e.* whether arterioles, capillaries, or veins are affected).

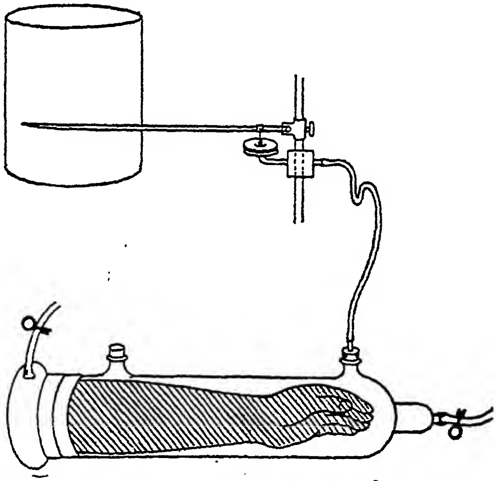


FIG. 179.—Human Plethysmography.

The forearm is enclosed in an airtight plethysmograph which is connected with a Marey's tambour. If a blood pressure cuff is placed on the upper arm and inflated for 10″ to arrest the venous outflow, the blood flow can be measured by venous occlusion plethysmography.

(2) MEASUREMENT OF ARTERIAL INFLOW BY VENOUS OCCLUSION PLETHYSMOGRAPHY.[1]—The organ is enclosed in a plethysmograph as above; the venous outflow is obstructed for a few seconds. The arterial inflow continues practically unchecked during this time and the organ swells to an extent which is recorded by the plethysmograph system. When the blood flow is to be measured by this method in a human limb (forearm, hand) a cuff is wrapped round the arm above the plethysmograph and the pressure is raised suddenly to 50–70 mm. Hg (*i.e.* just below the diastolic pressure) to occlude the veins for 10 seconds. The degree of limb swelling which occurs during this time is determined (Fig. 200); *it equals the arterial inflow.* The pressure is then released and the procedure repeated at intervals of 5–10 seconds.

It should be noted that in man the upper arm or forearm consists of 85% muscle and 10% skin; in the hand, and especially in the digits, the muscles form a much smaller proportion of the total tissue mass. Changes in skeletal muscle blood flow in man are discussed on p. 432, and in skin ~~flow on p. 327~~.

(3) SKIN TEMPERATURE.—These measurements are carried out in man with a thermocouple connected with a sensitive galvanometer. A fall of skin temperature (when the external temperature is constant) represents diminished local blood flow (usually due to active *arteriolar* constriction); conversely, a rise of temperature represents *arteriolar* dilatation and greater blood flow. (For records, see Figs. 198, 211, 216, 217.)

(4) CEREBRAL BLOOD FLOW.—(i) *Estimates based on arterial-venous* (*A-V*)

[1] Lewis and Grant, *Heart*, 1925, *12*, 73. Grant *et al.*, *Clin. Sci.*, 1938, *3*, 119 157, 273, Wilkins *et al.*, *ibid.*, 403.

oxygen difference.—In man the oxygen content of arterial and jugular venous blood may be determined on samples obtained by direct puncture of the vessels. The average results obtained are as follows: oxygen capacity 20 c.c. (per 100 c.c.); arterial O_2 content 19 c.c.; jugular venous O_2 content, 12·8 c.c., arterial-venous oxygen difference 6·2 c.c. The venous blood leaving the brain is thus more extensively reduced than elsewhere in the resting body, *i.e.* about 65% saturated with oxygen instead of 75% as in mixed venous blood. If it is assumed that in any one *normal* person the oxygen consumption of the brain is constant, it follows that an increase in the A-V oxygen difference indicates a decreased brain blood flow: a decrease in the A-V difference indicates a rise in the blood flow. In *abnormal* brain states (*e.g.* epilepsy) these assumptions may not hold good.

(ii) *Nitrous Oxide Method.*[1]—The subject breathes a mixture of 15% nitrous oxide in air for 10 minutes. During this period six samples of blood are taken at intervals from an artery and from the jugular venous bulb, and their nitrous oxide content is determined. After making a number of assumptions and employing a complex formula a value is obtained which is claimed to be the cerebral blood flow. The average flow in normal people is 54 c.c. (±12) per 100 g. of brain per minute. If the brain weight is taken as 1400 g. the total cerebral blood flow is 750 c.c. (range 600–900) per minute, or one-seventh of the total cardiac output. As the arterio-venous oxygen difference is about 6·5 c.c.%, the total oxygen consumption of the brain is about 50 c.c. per minute or over 3·5 c.c./100 g./min. The cerebral blood flow should be compared with that of muscle, which at rest is 1–4 c.c. and during violent exercise 30–40 c.c. per 100 g. per minute; the cerebral blood flow per 100 g. thus substantially exceeds that of powerfully contracting muscle. The metabolic rate of nerve cells must be extraordinarily high.

The cerebral blood flow is *diminished* by the following procedures:

(i) A fall of arterial blood pressure reflexly produced by carotid sinus stimulation.

(ii) Overventilation which lowers the arterial CO_2 tension; lack of CO_2 acts *directly* on the cerebral arterioles causing active constriction.

(iii) A rise of intracranial pressure above 300 mm. H_2O which mechanically compresses the cerebral capillaries and veins.

(iv) Inhalation of 85–100% oxygen which reduces the flow by 15%.

A *rise* in cerebral blood flow is produced:

(i) By CO_2 excess which acts directly on the cerebral vessels causing active dilatation. Inhalation of 5–7% CO_2 increases the cerebral blood flow by 75%.

(ii) By oxygen lack: inhalation of 10% O_2 increases the flow by 35%.

The cerebral and pial vessels are supplied by *sympathetic vasoconstrictor nerves.* Direct observation of the exposed pial vessels in animals shows that stimulation of the peripheral end of the cervical sympathetic causes constriction of the pial vessels, with or without a rise of blood pressure or of cerebrospinal fluid pressure; the average reduction in the diameter observed is 7%. The vessels are moderately dilated after sympathectomy, or reflexly by stimulating the central end of the vagus. The purpose served by the sympathetic innervation is not clear. The blood flow through the brain ***varies directly with the level of arterial blood pressure*** (*e.g.* it is markedly reduced

[1] Kety and Schmidt, *J. clin. Investig.*, 1948, *27*, 476 *et seq.*

in a severe hæmorrhage). As the skull cavity is a closed inexpansible box, dilatation of the cerebral vessels can occur only to the extent that cerebrospinal fluid is expelled; such dilatation results from local changes in the chemical composition of the blood, *e.g.* CO_2 excess, O_2 lack. The blood vessels are directly constricted by CO_2 lack or high O_2 pressure. A raised venous pressure mechanically impedes the outflow of blood from the brain. In disease, ***spasm*** of the cerebral vessels may occur from nervous or chemical causes.

(5) Coronary Circulation, see p. 237.

(6) Pulmonary Circulation.[1]—The pulmonary arterioles are few and contain little smooth muscle; the corresponding capillaries and veins are very distensible. There is thus a low peripheral resistance in the pulmonary circulation; the blood pressure in the right ventricle and pulmonary artery is correspondingly much lower than in the left ventricle and aorta. In man right ventricular pressure is 18–30 mm. Hg systolic (average 25) and 0–2 mm. diastolic; pulmonary artery pressure is on an average 25 mm. systolic and 8 mm. diastolic (Fig. 150); pulmonary capillary blood pressure, therefore, lies between 8 and 2 mm. Hg, which is far less than the plasma protein osmotic pressure (p. 15). The wall of the right ventricle is much thinner than that of the left ventricle.

Sympathetic fibres supply the media of the pulmonary arterioles; sympathetic stimulation causes constriction of these vessels, raising pulmonary artery pressure. The physiological significance of this nervous control is unknown.

The time taken by a red cell to flow through the lungs at rest is 4–5 sec.; the time spent by a red cell in the capillaries is 0·7 sec. at rest and 0·3 sec. during exercise; during this brief period gaseous pressure equilibrium is established between the alveolar air and the pulmonary capillary blood. The volume of blood present in the lungs at any time is normally about 10% of the total blood volume, *e.g.* 500 c.c. As the pulmonary vessels are so readily distended, they can accommodate larger volumes of blood without a significant rise of local blood pressure. An increase in pulmonary blood volume always takes place at the expense of the air in the air-sacs, thus decreasing the vital capacity (p. 362). For example, on passing from the standing or sitting to the recumbent position the venous return and right heart output are increased, causing " active congestion " of the lungs and a corresponding decrease in the vital capacity.

Pulmonary Circulation in Disease.—(i) If one lung is removed or completely *collapsed* by an artificial pneumothorax, there is usually no rise in pulmonary artery pressure because the vessels in the normal lung become sufficiently distended to transmit the larger blood flow through them with little increase in peripheral resistance.

(ii) In *emphysema* overdistension of the alveoli tends to compress and obstruct the vessels lying in their walls; similar effects may be produced by extensive pulmonary *fibrosis*. In advanced cases, as a result of the increased resistance of the blood flow, the pulmonary artery pressure is raised (*e.g.* to 45 mm. Hg systolic) and the work of the right heart is correspondingly increased.

(iii) In *congestive heart failure* (*e.g.* resulting from mitral stenosis) the left auricular pressure, and therefore pulmonary venous pressure, rises,

[1] Cournand, *Bull. N.Y. Acad. Med.*, 1947, *23*, 27.

Finally the right ventricle may fail to empty itself completely and right ventricular diastolic, as well as systolic, pressures rise considerably (Fig. 150).

Factors concerned in the development of *pulmonary œdema* are considered on p. 116.

(7) Splanchnic Area and Skin.—These two regions constitute the *major part of the peripheral resistance.* Stimulation of the peripheral end of the splanchnic nerve produces a considerable rise of blood pressure (*e.g.* as much as 100 mm. Hg) owing to constriction of the vessels in the viscera. The rise of pressure is usually partly due to the simultaneous secretion of adrenaline (cf. p. 731). Conversely, splanchnic section causes marked vasodilatation in the viscera (Fig. 187, A, p. 315).

The changes in the circulation in the human skin under various conditions are fully discussed on pp. 321 *et seq.*

(8) Renal Circulation.—This is fully discussed on pp. 24 *et seq.*

(9) Skeletal Muscle.—These arterioles receive both a vasoconstrictor and a vasodilator sympathetic innervation. Normally the tonic constrictor influence predominates; thus blocking the sympathetic supply to the muscle vessels in man, increases the resting blood flow to a moderate extent, *e.g.* from 3 or 4 to about 8 c.c. per 100 c.c. of tissue. On heating the body not only is constrictor tone in the muscle vessels inhibited, but the vasodilators are also reflexly stimulated producing a greater blood flow than occurs after sympathetic nerve block (which puts both vasoconstrictors and vasodilators out of action). It is noteworthy that when fainting occurs as a result of a hæmorrhage, the blood flow through the skeletal muscles is greatly increased in spite of the associated low level of arterial blood pressure; sudden vasodilatation due to nervous impulses seems to be the causal factor (p. 84). Minute doses of adrenaline dilate muscle vessels; larger doses may be constrictor.

The principal factors controlling the calibre of the skeletal muscle vessels are *metabolites, i.e.* the chemical products of muscular activity, or the substances released during ischæmia (p. 330).

(10) Large Peripheral Vessels.—The descending aorta, the iliacs, and the subclavians are innervated from adjacent sympathetic ganglia. The more peripheral vessels (axillary, femoral, etc.) are supplied *at intervals* by non-medullated postganglionic fibres which have travelled in adjacent nerve trunks and leave them to supply short lengths of the vessel wall.

(11) Innervation and Functions of Veins, p. 318.

(12) Innervation and Functions of Capillaries, pp. 17, 319.

(13) Determination of the *circulation time* may throw light on the state of the peripheral circulation, both venous and arterial (p. 262).

Vasodilator Nerves.—Vasodilator fibres are found as follows:

(1) *In the sympathetic nerves* [1]: the sympathetic fibres to blood vessels are, of course, predominantly vasoconstrictor in character, but by special methods a certain number of dilator fibres can also be demonstrated, *e.g.* supplying the muscle blood vessels in man (*supra*).

(2) *In the cranial* and *sacral autonomic*: *e.g.* chorda tympani (p. 20), ninth nerve (p. 714), nervi erigentes (p. 712.)

(3) *Vasodilator Fibres in the Dorsal Nerve Roots.*—As is explained on p. 721, stimulation of the peripheral end of the cut dorsal nerve root produces

[1] Burn, *Physiol. Rev.*, 1938, *18*, 137.

vasodilatation in the corresponding segment of the body. Impulses pass to the skin causing a local release of histamine which produces its characteristic triple response of capillary dilatation, wheal, and arteriolar dilatation (flare) (p. 325) (Fig. 195). Other impulses pass to the muscle vessels releasing acetylcholine which causes vasodilatation.

Herpes Zoster.—The above experiments explain the cutaneous phenomena of *herpes zoster.* In this disease noxious influences, such as inflammation, hæmorrhage, pressure, or poisons, act on the dorsal root ganglia or their cranial homologues. The lesions are equivalent to stimulation of the ganglia, and so dilatation and increased permeability of the blood vessels in the corresponding skin area occur. This leads to exudation of protein-rich fluid and formation of *vesicles* in groups along the distribution of the dorsal nerve roots or their cutaneous sensory branches.

Factors influencing Vasomotor Centre.—(1) Higher Centres.—(i) A rise of blood pressure commonly occurs in emotional stress and in

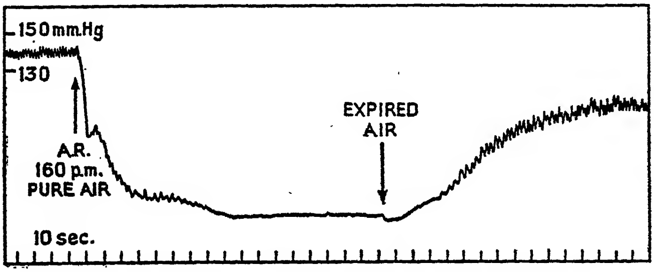


Fig. 180.—Effects of Overventilation on Blood Pressure.

Cat. Blood pressure. At the first arrow, the animal was ventilated with pure air at the rate of 160 times per minute. The blood pressure rapidly falls. At the second arrow, expired air (rich in CO_2) was substituted. The blood pressure rises again. (Dale and Evans, *J. Physiol.*)

anticipation of muscular exercise; (ii) inhibition of the vasomotor centre may in part be responsible for fainting fits, *e.g.* in anæmic girls (*vaso-vagal attack*) (p. 271). Both effects are due to impulses from the higher centres to the vasomotor centre.

The effects of stimulation of the *hypothalamus* and *cerebral cortex* are discussed on pp. 716, 671.

(2) Radiation from Respiratory Centre.—During inspiration, impulses flow from the respiratory centre setting up increased activity of the vasomotor centre (pp. 303, 313); during expiration the discharge of the centre is diminished.

(3) CO_2 Tension.—The vasomotor centre can only function effectively in the presence of an *adequate CO_2 tension.* If an animal's lungs are excessively ventilated with atmospheric air to lower the CO_2 tension in the alveolar air and consequently in the arterial blood, a considerable fall of blood pressure usually results; this is principally due to inhibition of the vasomotor centre resulting in dilatation of the vessels in the splanchnic area (Fig. 180). The blood pressure can be restored and raised above normal by ventilating with air containing excess of CO_2, thus restoring the CO_2 pressure in the arterial

blood. The inhibitory and stimulating actions of CO_2 on the vasomotor centre are partly direct; in part, however, they are *reflexly* produced via the sino-aortic nerves from the chemoreceptors in the carotid and aortic bodies (p. 745).

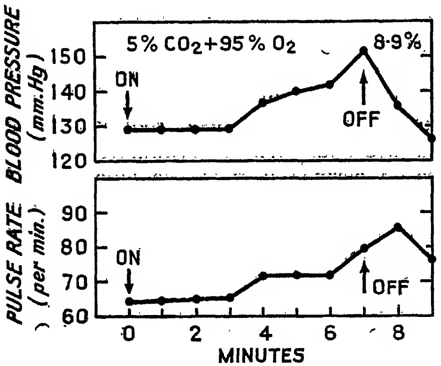


FIG. 181.—Effect of CO_2 Excess on Human Blood Pressure and Pulse Rate.

At ON, subject rebreathed from a bag containing 15 L of a mixture of 95% O_2 and 5% CO_2. At OFF, the subject was again put on to air; the final value of CO_2 in the bag was 8·9%. (Wright.)

By means of perfusion experiments it can be shown that CO_2 also acts *directly* on the arteriolar wall: a *raised CO_2 tension relaxes arterioles*, a lowered CO_2 tension constricts arterioles (cf. p. 317). The peripheral effects are very evident in the *skin*. By its *peripheral action*, therefore, CO_2 produces effects which are *just the opposite of those it produces by its central action.*

In *man* the central and peripheral effects of CO_2 may neutralize one another to a considerable extent. *Overventilation* in man (which lowers the CO_2 tension) commonly produces no change in blood pressure; the skin vessels are *constricted*, and the skin becomes cold and pale (p. 408). CO_2 excess in man only raises the blood pressure after a considerable latent period and when the CO_2 pressure is very high, *e.g.* 7% of an atmosphere (Fig. 181). The skin becomes red and hot, *i.e.* vasodilatation takes place there, though vasoconstriction presumably occurs in the internal regions.

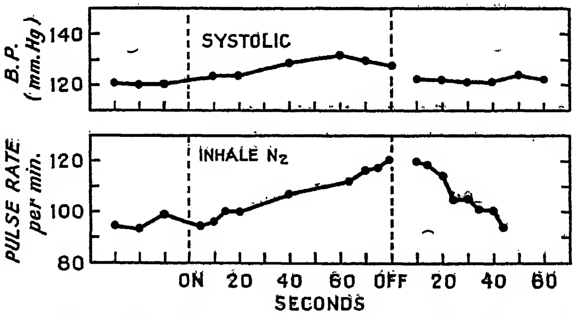


FIG. 182.—Effect of Acute Anoxia on Human Blood Pressure and Pulse Rate.

To left of first vertical line=control readings. Between ON and OFF nitrogen was inhaled (cf. Fig. 243). (After Schneider and Truesdell, *Amer. J. Physiol.*)

(4) OXYGEN LACK.—O_2 *excess* has no effect on the vasomotor centre. A severe *decrease* in the oxygen tension of the blood—produced in animals by artificially ventilating the lungs with an inert gas like nitrogen, or by giving an oxygen-poor mixture—stimulates the vasomotor centre (partly directly,

mainly reflexly via the sino-aortic nerves from the chemoreceptors), and may cause a considerable rise of blood pressure. It must be remembered, however, that anoxia depresses the heart's action: if the heart weakens markedly, no rise, or even a fall of pressure may occur.

In *man* acute severe anoxia produces little rise of blood pressure, and most of that is emotional in origin (Fig. 182). After about one minute, with the onset of cyanosis, the pressure may fall, sometimes considerably, owing to the heart muscle becoming depressed. Anoxia which develops *gradually* in man may lead to no change in arterial blood pressure before the onset of heart failure (Fig. 243).

Cerebral Anæmia.—Acute complete *cerebral anæmia* (from occlusion of the carotid and vertebral arteries) may produce a huge rise of blood pressure in animals, partly reflexly from the collapsed carotid sinuses and partly directly because of CO_2 accumulation and oxygen lack in the brain (Fig. 183).[1] Some rise of blood pressure may occur clinically from cerebral anæmia due to raised intracranial pressure.

Changes in the *pressure in the cerebral vessels* have *no* direct effect on the tone of the vasomotor centre.

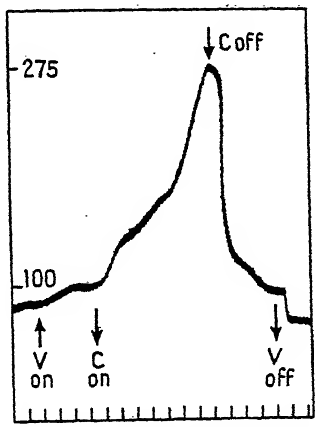


FIG. 183.—Effect of Cerebral Anæmia on Blood Pressure.

Ordinate: Arterial blood pressure in mm. Hg. Time in minutes.
Sinus nerves intact. Vagi cut. At first arrow (*V on*) clamp vertebral arteries; at second arrow (*C on*) occlude both carotids. The immense rise of pressure which results is partly reflex and partly due to stimulation of the vasomotor centre by *complete anæmia*. At *C off*, release carotids and at *V off*, release vertebral arteries. (Wright, *J. Physiol.*, 1930.

(5) SINUS AND AORTIC NERVES.—The afferents from the *pressure receptors* are principally responsible for *stabilizing the blood pressure* and preventing it from varying very much; they similarly maintain the resting heart rate within narrow limits. These sino-aortic nerves are normally carrying up a constant stream of afferent impulses which exert a tonic *inhibitory* influence on the vasomotor centre (cf. their action on heart rate, p. 272).[1] The evidence is set out elsewhere (p. 738); The main points only are referred to below.

(i) Stimulation of the central end of the aortic or sinus nerves usually reflexly lowers the blood pressure, in part owing to inhibition of the vasomotor centre with resulting vasodilatation.

(ii) *Section* of either pair of nerves raises the blood pressure (more especially if the other pair has previously been cut) mainly because of released overactivity of the vasomotor centre and resulting vasoconstriction. Denervation in stages leads to the development of *chronic hypertension* (p. 744).

(iii) Variations in carotid sinus pressure can be produced by occluding the carotid arteries or by appropriately planned perfusion experiments. A fall of carotid sinus pressure reflexly produces vasoconstriction and a rise of blood pressure, while a rise of carotid sinus pressure reflexly produces vasodilatation and a fall of blood pressure.

[1] The afferent impulses from the *chemoreceptors* set up by O_2 lack or CO_2 excess also eflexly *stimulate* the vasomotor centre (see pp. 743, 746).

(iv) The high sensitivity of the sinus to internal pressure changes must be stressed ; a deviation of 10 mm. Hg is sufficient to cause a reflex response. Koch studied the reflex effect on blood pressure of raising the pressure in the isolated innervated carotid sinus from zero to various levels.[1] The results are plotted, with the pressure in the sinus along the abscissa and the percentage fall of blood pressure on the ordinate (Fig. 184). An S-shaped curve is obtained. The main features are these : there is a threshold value for the sinus pressure below which no reflex is set up (this varies from 30 to 70 mm. Hg) ; at the other extreme there is a maximum effective pressure beyond which no additional response is obtained (this varies between 150 and 300 mm. Hg). The curve is steepest in the normal blood pressure range, and it is interesting to note that the actual turning-point, marked with — on the tracing, agrees closely with the normal average level of blood pressure in the

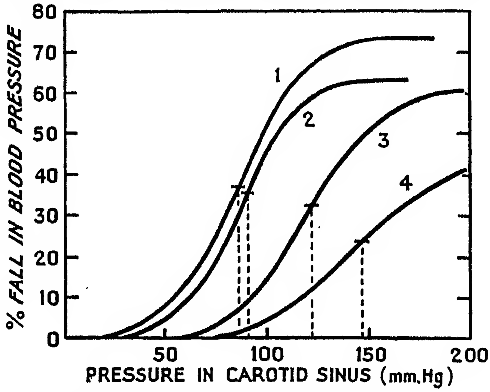


FIG. 184.—Carotid Sinus Pressure and Vascular Tone.

The curves show the percentage fall of blood pressure (ordinate) produced by raising the pressure in the isolated carotid sinus (abscissa). Up to a point the higher the level of carotid sinus pressure the greater the reflex fall produced. The point marked with a — is the level of normal blood pressure for the species and is the region in which the most sensitive sinus reflexes are obtained. At higher sinus pressures progressively smaller reflex effects are obtained. 1=monkey ; 2=rabbit; 3=dog; 4=cat. (After Koch.)

particular species studied : the value for the monkey is 85–95, the dog 110–135, and for the cat 145 mm. Hg. Variations in sinus pressure on either side of this point cause the greatest reactions. In the intact animal, therefore, when the blood pressure tends to deviate from the normal, the receptors in the sinuses respond in a very sensitive manner, affect the bulbar centres appropriately, and rapidly restore the blood pressure to its usual level.

The afferents arising in the aortic arch presumably act similarly.

(6) AFFERENTS FROM OTHER REGIONS.—General or local vasomotor tone can be reflexly modified by afferent impulses from various parts of the body.

(i) In animal experiments, stimulation of the *central* end of a peripheral nerve (*e.g.* sciatic) with strong, rapidly interrupted currents reflexly raises the blood pressure ; through a certain range of stimuli, the greater the strength of stimulus the greater the rise of pressure. Weak or slowly inter-

[1] The other sinus nerve and both vagi were cut to exclude cardiac and compensatory effects.

rupted currents produce a fall.[1] The peripheral nerves thus contain both afferent *pressor* and afferent *depressor* fibres which respectively stimulate and inhibit the vasomotor centre.

(ii) In man, a variety of vasomotor reflexes can be obtained. Thus (*a*) a sensory stimulus like a noise may reflexly constrict skin vessels and dilate muscle vessels (Fig. 185) ; (*b*) heat and cold applied to the skin respectively reflexly dilate and constrict the skin vessels, especially in the hands and feet (p. 327) ; (*c*) distension of the duodenum reflexly constricts the vessels in the toes and fingers (p. 762).

Relation of Respiration to Blood Pressure.—In animal experiments it is common to find regular variations in the blood pressure with the phases of respiration, the pressure rising during inspiration and falling with expiration ; a similar relationship may sometimes be noted in man (Fig. 186). On the other hand, the opposite relationship may be found, or the blood

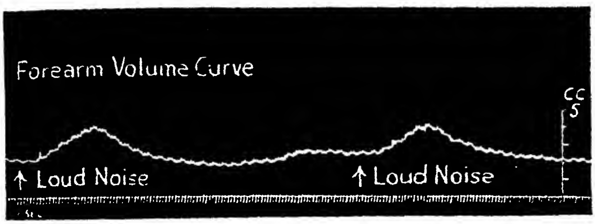


FIG 185.—Reflex Dilatation of Muscle Vessels in Man. (Grant and Pearson, *Clin. Sci.*, 1938, *3*, 135.)

Forearm (volume = 600 c.c.) in plethysmograph ; a loud noise (at the arrows) reflexly increases forearm volume (vasodilatation). As the hand under the same conditions shows vasoconstriction it may be supposed that the skin vessels constrict ; as the forearm volume increases in spite of cutaneous vasoconstriction, the muscle vessels have dilated.

pressure may be unaffected by the phases of respiration. The reason for this variability is that conflicting factors are at work.

During *inspiration* there is (i) irradiation of impulses from the respiratory centre to the vasomotor centre, increasing arteriolar tone (cf. p. 303) ; a deep inspiration in man, for example, produces notable constriction of the skin vessels (*e.g.* in the fingers) ; (ii) owing to the descent of the diaphragm and the greater negative pressure in the pleural cavity the venous return is increased, and the heart rate and output rise similarly. These two factors tend to raise the blood pressure. On the other hand, during inspiration (iii) the negative pressure in the thorax is increased. This pulls on the walls of the thin capillaries and venules which surround the alveoli of the lung and increases the lumen of these vessels. As a result, some of the blood which is expelled by the right heart is held up in the lungs to fill the dilated vessels, and less therefore reaches the left heart. The left ventricular output is lessened, and this tends to lower the systemic blood pressure.

During *expiration* the intrathoracic negative pressure is diminished ; the result is that some of the excess blood is squeezed out from the pulmonary bed. The left heart receives this blood in addition to the output of the right heart, tending to raise the blood pressure ; but owing to the decrease in

[1] This rule does not apply to the afferent fibres in the vagus nerves (Fig. 475).

general venous return during expiration, less blood reaches the heart, the output is decreased, and the blood pressure tends to fall.

It will be noted that in the main two conflicting factors are constantly at work: the changes in the capacity of the pulmonary bed tend to lower blood pressure in inspiration and raise it in expiration, while the venous return and heart rate alterations tend to produce exactly the reverse effects. It is not surprising, therefore, to find that the effects of respiration vary in different species and with the exact type of breathing employed.

Relation of Adrenal Medulla to Blood Pressure.—This is fully considered on pp. 724, 731 *et seq.*

There is no evidence that any of the other ductless glands, *e.g.* the neural division of the pituitary (p. 46), play any part in normal blood pressure control.

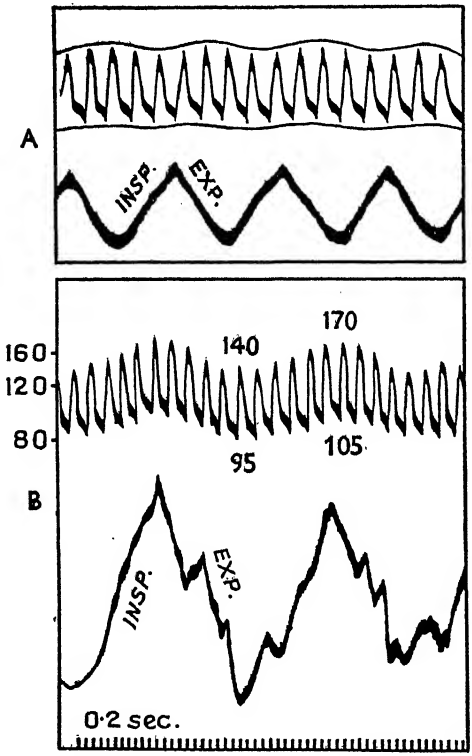


FIG. 186.—Variations in Blood Pressure with Phases of Respiration in Man. (Battro *et al.*, *Arch. int. Med.*, 1944, *73*, 29.)

A: quiet breathing. B: deep breathing.

In both A and B the arterial blood pressure was recorded by means of an *intra-arterial* cannula connected with a membrane manometer. Time in B in 0·2 sec.

B.P. values in B are in mm. Hg. In A the blood pressure is maximal at the end of (quiet) expiration and minimal at the end of (quiet) inspiration. In B with deep breathing the fluctuations are more marked; B.P. is maximal at the end of inspiration and minimal at the end of expiration; thus systolic B.P. varies between 140 and 170 mm., and diastolic between 95 and 105.

Regulation of the Normal Blood Pressure.—In the normal resting person the blood pressure is steadily maintained within comparatively narrow limits. If one may so express it the normal level of pressure represents a compromise between a number of conflicting interests. It must be remembered that, other things being equal, the blood flow through the brain will vary directly with the blood pressure. A pressure which is much below normal is unable to drive an adequate amount of blood to the brain against the force of gravity. If the blood pressure is raised excessively, it serves no useful purpose and imposes an additional burden on the heart and a greater strain on the blood vessel walls. The level of blood pressure also regulates the coronary blood flow (p. 237) and the rate of filtration in the glomeruli of the kidney (p. 25). The desired *constancy* of the blood pressure is attained mainly reflexly by the sino-aortic nerves (*buffer nerves*), by means of the variable tonic inhibitory influence which the *pressure-receptor* fibres exert on the

cardiac and vasomotor centres and on adrenaline secretion. A rise of blood pressure at rest increases their inhibitory activity, a fall of blood pressure lessens it.

The four sino-aortic nerves form a functioning entity; their regulatory rôle in the body can be illustrated by many observations:

(i) Occlusion of the carotid arteries produces a greater reflex rise of blood pressure after the aortic nerves have been cut and their restraining influence thus removed.

(ii) The rise of blood pressure resulting from stimulation of the peripheral end of the splanchnic nerves is greater after cutting the sino-aortic nerves, although the initial blood pressure is higher.

(iii) Stimulation of the central end of the aortic nerve in the rabbit

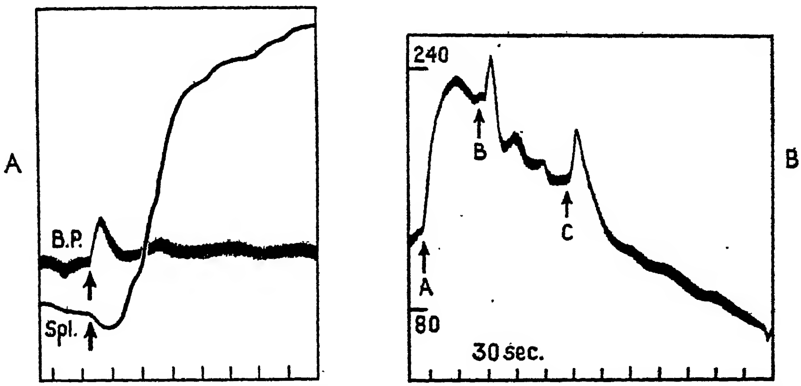


FIG. 187.—Effects of Splanchnic Nerve Section on Circulation. (Kremer and Wright, *Quart. J. exp. Physiol.*, 1932.)

A. Cat: Upper record, blood pressure. Lower record, spleen volume. Time tracing shows 30-second intervals. Aortic and sinus nerves intact. At the point marked by the arrows the right splanchnic nerve was divided. There is a slight initial rise of blood pressure and splenic constriction owing to mechanical stimulation of the nerve. This is followed by marked dilatation of the spleen from loss of vasoconstrictor tone. The blood pressure is unaltered owing to the compensatory effects set up reflexly by the pressure-receptor afferents in the sinus and aortic nerves.

B. Cat: Blood pressure record. *Both vagi cut. Right carotid clipped.* At the first arrow (A) left carotid occluded. Owing to the fall of sinus pressure the B.P. rises markedly to 240 mm. Both carotid sinuses are now out of action and the aortic nerves are, of course, divided. At the second and third arrows (B, C) the right and left splanchnic nerves are successively divided. After a transient initial rise (due to mechanical stimulation) the blood pressure rapidly and progressively declines.

brings about a bigger fall of blood pressure and a more sustained slowing of the heart when the other buffer nerves have been divided; the same is true of stimulation or distension of the carotid sinus.

(iv) If the splanchnic nerves are cut, with the sino-aortic nerves intact, the fall of blood pressure, if it occurs at all, is slight, transient, and soon recovered from in spite of extensive dilatation of the abdominal vessels (Fig. 187, A); this is due to reflex compensatory vasoconstriction in other parts of the body. If the sino-aortic nerves have been previously divided, splanchnic section produces a larger absolute and percentage fall of blood pressure (Fig. 187, B).

After hæmorrhage it is found that section of the vagi may produce a further fall of blood pressure as shown in Fig. 188. Such results were thought to indicate that in states of extreme hypotension afferent *pressor* impulses

(2) By *general circulatory adjustments.*—If the blood pressure is raised as a result of vasoconstriction limited to certain areas, blood vessels in other zones become passively opened up by the higher pressure of the blood reaching them. Thus injection of adrenaline (p. 724) constricts the vessels in the splanchnic area and skin and raises the blood pressure; the blood is thus diverted to the skeletal muscles and heart (where adrenaline also produces active dilatation of the vessels). The same type of adjustment occurs in muscular exercise, except that here in addition products of metabolism widely dilate the vessels in the skeletal muscles (p. 432).

(3) By *inhibition of vasoconstrictor tone,* usually reflexly via the vasomotor centre.—Thus heating the body dilates the blood vessels in the limbs by inhibiting the discharge of the vasomotor centre (p. 327).

(4) By *heat* acting *directly* on the vessel walls.—The skin vessels are thus opened up in hot weather; the temperature of active organs rises with a similar result (p. 327).

(5) By *chemical changes in the tissues.*—Perfusion experiments show that a *rise of* CO_2 *tension* or of H^+ *ion concentration* or *oxygen lack* directly dilates blood vessels. These agents may all come into action in secreting glands or contracting muscle. If muscle is made to contract anærobically after poisoning with iodoacetic acid, though no CO_2 or H^+ ions are liberated vasodilatation occurs; this observation shows that *non-acid* dilator substances may be liberated during activity. Experiments on *ischæmia* followed by *reactive hyperæmia* (p. 329) prove that when the blood supply to a tissue is cut off, the blood vessels locally dilate owing to the accumulation of certain products of tissue activity; *adenylic acid* and possibly *histamine* are among those chemically identified. These results suggest that a variety of products of tissue metabolism, called non-committally *metabolites,* are constantly formed at rest and appear in increased amounts as a result of activity and adjust (by variations in their concentration) the calibre of the local blood vessels, and consequently the local blood flow, at a level which is appropriate to the physiological state prevailing at the time.

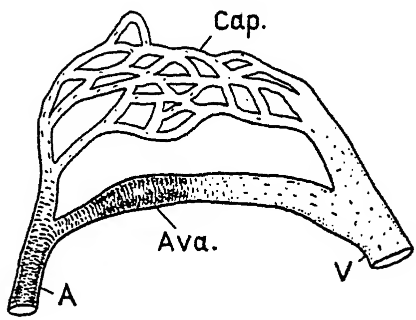


FIG. 189.—Diagram showing an Arterio-Venous Anastomosis (Ava) linking up a Terminal Arteriole (A) with a Small Venule (V) in the skin. Above is indicated the Capillary bed (Cap). (Le Gros Clark, *Tissues of the Body,* Oxford, 1939.)

These principles are discussed in greater detail for salivary glands (p. 20), heart (p. 238), skeletal muscles in exercise (p. 432), and the skin during exposure to varying external temperatures (pp. 327 *et seq.*).

Arterio-Venous Anastomoses.[1]—In many tissues the arterioles and venules are directly connected by channels which are independent of the capillary network (Fig. 189); the muscular wall of these cross-connections is very thick and richly supplied with vasomotor nerves. The anastomoses

[1] Clark, *Physiol. Rev.*, 1938, *18*, 229.

are particularly numerous in the skin, especially in the fingers in the region of the nail bed: they are opened up by heat, thus enabling a larger volume of blood to pass through these wide channels directly from the arteries to the veins; the resulting increase in dermal blood flow facilitates heat loss. The reverse changes are produced by cold. The anastomoses probably account for the enormous changes of skin flow produced by alterations of temperature (Fig. 198, p. 327). Anatomoses are also found in the sex organs, *e.g.* between the arteries and the cavernous sinuses of the penis, and in the kidney, liver, spleen, and lungs. The maximum diameter of the anastomoses in different species is 50–390 μ.

Pathological arterio-venous anastomoses are established when an injury affects the walls of an adjacent artery and vein leading to a direct blood flow from the artery to the vein (p. 298).

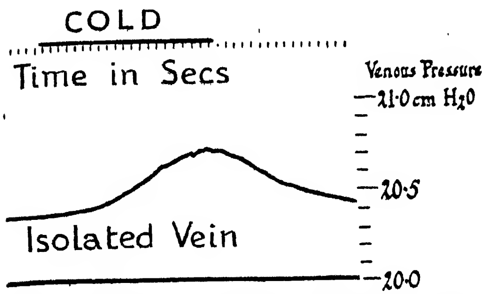


FIG. 190.—Reflex Contraction of Veins in Man (Doupe *et al.*, *J. Physiol.*, 1938, *92*, 396.)

An 8-cm. length of a wrist vein in man was exposed, blocked above and below, and needled; the contained blood was replaced by saline and the needle connected with a pressure-recording apparatus. During the period indicated by the signal line, cold was applied to the surface of the body. The vein *reflexly contracted* as shown by the rise of intravenous pressure.

Physiology of Veins.[1]—The veins are innervated and their calibre regulated in the same way as the arterioles. The sympathetic exerts a tonic constrictor influence on veins; sympathetic section leads to venodilatation in man. The veins constrict during vascular pressor reflexes and dilate during depressor reflexes; Fig. 190 illustrates *reflex* contraction of a human vein in response to cold. The veins are constricted when the CO_2 tension in the medulla is raised. As the veins contain so little muscle tissue the changes in calibre which they undergo under these conditions are small.

VENOUS PRESSURE.[2]—To estimate the venous pressure in man two main methods—the indirect and the direct—have been employed. The *indirect* methods consist in determining the external pressure necessary to produce collapse of a small superficial vein, or the pressure which just prevents refilling of a vein previously emptied by " milking " towards the heart. The *direct* methods consist in placing a manometer in communication with a vein by means of rubber tubing and a hollow needle of adequate bore. The vein usually chosen is the median basilic, at the bend of the elbow.

(i) Normal values are 5–15 cm. H_2O. Venous pressure is, of course, lower than capillary pressure, but the difference is small.

(ii) Venous pressure does not necessarily follow changes in the pressure in the main arteries; but if the arterioles are dilated locally, the general blood pressure is transmitted more completely to the veins and the pressure within them rises; this is seen in the submaxillary veins during secretory activity.

(iii) The venous pressure is increased during exercise because the raised

[1] Franklin, *Veins*, Springfield, 1937. Burch, *Primer of Venous Pressure*, 1950.
[2] Bedford and Wright, *Lancet*, 1924, ii, 106.

arterial pressure is transmitted through the very dilated arterioles of the active muscles.

(iv) Excessive hyperpnœa may lower the venous pressure slightly.

(v) A forced expiration with the glottis closed raises the venous pressure considerably, because the raised intrathoracic pressure compresses the venæ cavæ and this prevents the venous return into the chest.

(vi) When congestive heart failure occurs the venous pressure rises, but not before obvious clinical evidence of venous engorgement is present, *e.g.* enlargement of the liver and distension of the cervical veins. In marked congestive failure the basilic vein pressure may be as high as 30 or 35 cm. H_2O (p. 111). The height of the venous pressure can be roughly determined by noting the height of the column of blood in the neck veins above heart level.

The *control of the venous return* is discussed on p. 274.

THE CAPILLARIES. VASCULAR RESPONSES OF THE HUMAN SKIN[1]

The Capillaries.—The capillaries are lined by thin, flattened, nucleated polygonal cells joined together by cement substance. Although the capillaries have no muscle coat they can actively modify their own calibre and respond to nervous, hormonal and other chemical and physical stimuli. The active variations in capillary calibre are attributed to alterations in the state of the lining endothelial cells themselves; when these swell the capillary lumen is reduced.

With the *skin microscope* the capillaries of the human skin can be studied, *e.g.* at the root of the nail, where they run parallel to the surface, and a considerable length of loop may be examined. Similar direct studies can be made in the frog in the mesentery or web of the foot.

Independent Contractility of the Capillaries.—The capillaries can alter their calibre independently of the state of the arterioles or venules. When their tone is high they can resist the distending force of a raised venous or arteriolar pressure; but when their tone is low they tend to follow passively changes in the arteries and veins. This can be shown by direct observation in man and lower animals; it is also supported by much indirect evidence:

(i) From skin colour and temperature (p. 321).
(ii) From the action of adrenaline and post-pituitary extracts (p. 322).
(iii) From the triple response in human skin (p. 322).
(iv) From the effects of injecting histamine in the cat (p. 337).
(v) From the white line produced by lightly stroking human skin (p. 321).

Frequently, however, the calibre of the capillaries changes actively in the same direction as that of the arterioles and veins (*e.g.* when heat is applied to a part, p. 327).

Capillary Blood Pressure. Interchanges Between the Blood in the Capillaries and the Interstitial Fluid. Lymph.—These subject are fully considered on pp. 16–21.

[1] Lewis, *Blood Vessels of the Human Skin and their Responses*, London, 1927. Krogh, *Anatomy and Physiology of the Capillaries*, New Haven, 2nd edn., 1929. Drinker and Yoffey, *Lymphatics, Lymph and Tissue Fluid*, Cambridge, Mass., 1941.

Importance of Capillary Tone.—Compared with the total number of capillaries present in a muscle, the number which normally contain blood is very small. Indian ink is injected into the vessels of a *resting* muscle, the tissue fixed, and sections cut and studied: very few capillaries can be seen on examination, and most of these are very narrow. When the experiment is repeated with *tetanized* muscle, ten or more times as many capillaries are visible in any unit area, and they are all widely dilated (Fig. 191). Now capillaries contain no valves, so that when blood comes to a part, it must first fill the capillaries before it can pass on into the veins. Once it reaches the veins its return to the heart is relatively easy. The contractions of the skeletal muscles drive the blood on, and it cannot pass back owing to the presence of the valves. In the capillaries, however, muscular activity merely squeezes the blood from one capillary to another, and the blood returns in the intervals between the contractions. If the capillaries *all* over the body relax (as after the injection of histamine, p. 337), they retain so much blood within them that after they have become filled very little blood is available to reach the veins. As a result the venous return falls, the output of the heart diminishes, and the brain and tissues generally suffer severely from oxygen lack owing to their impaired blood supply. Normal capillary tone is thus of extreme importance. In vigorous exercise the muscular capillaries, it is true, relax, and can retain a great deal more blood, but then the blood flow to the muscles is greatly increased partly by diminishing the lumen of the blood vessels in other regions of the body.

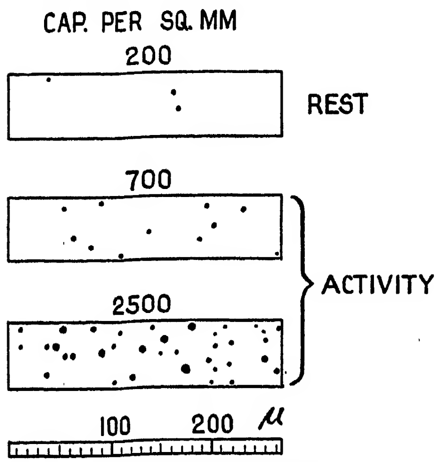


FIG. 191.—Effect of Activity on Capillary Tone.

Capillaries vitally injected with Indian ink. Transverse sections of muscle. Note increase in number and calibre of patent capillaries as a result of activity. Scale in μ. (Krogh, *Anatomy and Physiology of Capillaries*.)

Sympathetic fibres pass to the capillary wall and probably transmit tonic vasoconstrictor impulses from the vasomotor centre. In the main, however, capillary tone is controlled by *local physical and chemical agencies*.

Control by Physico-Chemical Agencies.—(1) METABOLITES.—During activity marked dilatation of the capillaries (and arterioles) takes place, as was noted above in the case of skeletal muscle. The dilatation is brought about by products of tissue activity, *i.e. metabolites*; the factors at work may perhaps be the rise of CO_2 tension or of H^+ ion concentration or the appearance of some non-acid dilator substances. When the blood supply to a limb is cut off, the locally acting physiological dilator substances accumulate and are responsible for the vasodilatation, affecting capillaries and arterioles in both skin and muscle, which becomes evident when the circulation is reopened (*reactive hyperæmia*, p. 329).

When human skin is injured, *histamine* is liberated and is responsible for the local vascular reaction. Other chemical agents may be released in damaged tissues elsewhere (*e.g.* in burns, p. 344).

(2) Both *adrenaline* and *post-pituitary extract* on injection produce local capillary constriction. Rigid proof of this in man is given on p. 322. If post-pituitary extract is absorbed (*e.g.* following intramuscular injection), it produces widespread capillary constriction especially of the face. (p. 46). There is no evidence that either hormone normally regulates capillary tone.

(3) For effects of *heat*, *cold*, and *light*, see pp. 327 *et seq*.

Vascular Responses of the Skin.—CONTROL OF SKIN COLOUR AND TEMPERATURE.—The *colour* of the skin depends on the calibre of the surface *capillaries*, and of the sub-papillary venous plexuses (*minute vessels*); when they are dilated and a considerable blood flow takes place through them, the colour is red; when the blood stagnates in relaxed capillaries, the oxyhæmoglobin undergoes extensive reduction, and so the skin is blue; when the capillaries are contracted, the skin is pale. The *temperature* of the skin, like that of any part, depends on the amount of blood *flowing* through it, which in turn depends on the calibre of the *arterioles*. The capillaries thus regulate the amount of blood which is present *in* a part at any moment, and the arterioles the amount passing *through* the region during any given period. In warm weather the skin is warm and the colour is pale or red. A warm red skin indicates that both the arterioles and the capillaries are relaxed. A warm pale skin must mean dilated arterioles and constricted capillaries. In cold weather the skin is cold and the colour is white, red, or blue. Cold pale skin means that both arterioles and capillaries are constricted. A cold blue or red skin means constricted arteries and dilated capillaries. The above observations prove that the capillaries in part can and do actively regulate their own calibre.

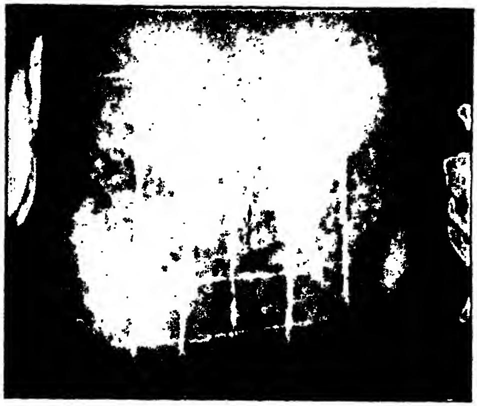


FIG. 192.—White Line produced by Light Stroking of the Skin. (Krogh, *Anatomy and Physiology of the Capillaries*.)

WHITE LINE.—If the skin on the front of the abdomen or forearm is stroked quite lightly, a *white line* normally appears which is exactly limited to the area of skin stimulated, and is quite sharply outlined (Fig. 192). The line appears after a latent period, persists for a minute or two, and then gradually fades. The white line is due to local active capillary contraction.[1] The evidence is as follows:

(i) Each arteriole supplies an *irregular* area of skin, so that the local anæmia which results from its contraction would similarly be irregular

[1] Cotton, Slade, and Lewis, *Heart*, 1917, 6, 227.

in outline and not follow strictly the line of the stroke as does the white line.

(ii) If the pressure in a cuff round the arm is rapidly raised to well above the arterial blood pressure (*i.e.* to 200 mm. Hg), the white line can still be obtained by stroking the skin of the forearm, and its time relations are similar to those in a part in which the circulation is unoccluded. The same holds true even if an interval of 10 minutes is allowed to elapse after applying the armlet. The effect of compressing the arm in this experiment is to stop the flow in the vessels distal to the block. In such a system, constriction of the *arterioles would drive blood out of them into the capillaries* and dilate the latter still further, as their contents cannot pass on in the normal way, owing to the obstruction to the natural outflow from the veins. The blanching of the skin, which is produced by stroking, must therefore be due to the contraction of the *capillary* wall, which actively drives the blood out into the neighbouring veins. This is confirmed by microscopic examination of the skin vessels.

FIG. 193.—Red Line produced by Scratching the Skin. (Krogh, *Anatomy and Physiology of the Capillaries*.)

(iii) The white line can be elicited after all the nerve fibres to the skin have degenerated, proving that it is a *direct local* response of the capillary wall to external stimulation.

Action of Hormones.—If *adrenaline* or *post-pituitary extract* is injected into the forearm skin after the circulation to the part has been arrested, it also causes local blanching, proving that these drugs produce active capillary contraction.[1]

Contractile Power and Resistance of Capillary Wall.—It can be shown that the maximum pressure which the capillary wall can exert on its contents when it contracts is 40–60 mm. Hg. If the pressure in a cuff surrounding the arm is raised to this figure, the venous pressure below the cuff rises to just above this level; the capillary pressure in the forearm skin must therefore be of a higher order (p. 17). If, under these conditions, adrenaline is injected into the forearm skin it still causes blanching. If the capillaries are *first* constricted by adrenaline, and the venous pressure is subsequently raised by obstructing the outflow of blood from the limb, they can resist a distending force in the veins of 80–100 mm. Hg; *i.e.* the resistance of a closed capillary is greater than the maximum pressure which it can exert on its contents when it contracts.

The Triple Response.—In subjects with more sensitive skins, or when heavier stroking is employed, a triple response may be obtained.

(1) RED LINE.—After a short latency a *red* line develops which it sharply defined (Fig. 193); it is first bright red in colour, but gradually

[1] Lewis, *Heart*, 1924, *11*, 109.

assumes a bluish tinge, indicating that the blood is stagnant and that its hæmoglobin is therefore undergoing excessive reduction. Using the criteria described above, the red line can be shown to be due to *capillary* dilatation ; it also can be produced after the local nerve supply has degenerated.[1]

FIG. 194.—Red Line and Flare in response to Scratching of Skin. (After L. R. Müller, from Krogh, *Anatomy and Physiology of Capillaries*.)

(2) FLARE.—In sensitive subjects, after a further interval of 15–30 seconds, a widespread flush or flare appears (Fig. 194) which is primarily the result of *arteriolar* dilatation :

(i) The flare spreads for a variable distance from the area stimulated, is mottled in colour from alternating islets of hyperæmia and of normal skin, and the edge is crenated, as would be expected from the manner in which the arterioles are distributed to the skin.

(ii) The colour is red throughout and does not tend to become purple from oxygen deficiency, proving that rapid flow of blood is occurring.

(iii) The temperature of the skin over the flare rises by 0·5–2° C. as a result of the increased blood flow.

The flare depends on an *intact nerve supply* to the skin ; it is abolished when the local nerve fibres have degenerated. It is still obtained if conduction in the sensory nerves leading from the part is blocked by an injection of procaine into the nerve trunk ; the reaction is therefore not a reflex through

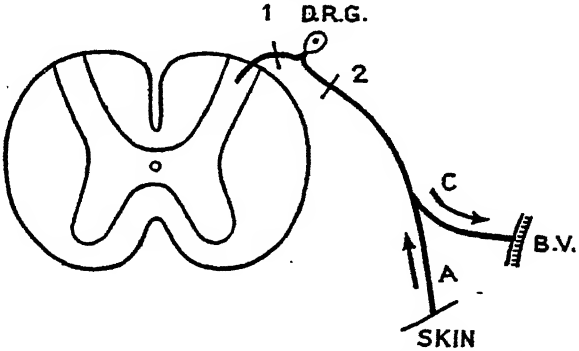


FIG. 195.—Antidromic Vasodilators in Dorsal Nerve Roots, Axon-Reflex in the Skin and Mechanism of Flare. (Cf. p. 721.)

D.R.G. = Dorsal root ganglion ; B.V. = Cutaneous blood vessel.

the central nervous system, but results from a *local axon reflex* (Fig. 195), *i.e.* a response which depends on the branching of nerve fibres in the skin. As will be shown later (p. 325) scratching the skin liberates histamine which stimulates local nerve endings ; impulses pass up the afferent nerve

[1] The red line is identical with the " tache cérébrale " found commonly in meningitis but long recognized as having no diagnostic significance.

A and along its branch C to dilate local arterioles giving rise to the "flare." Histamine injected into the skin acts similarly.

(3) WHEAL.—In susceptible subjects a wheal develops on the site of the red line (Fig. 196). It reaches its maximum height in about 5 minutes, and may project 1–2 mm. above the general surface of the skin. It is due to *increased permeability* of the minute vessels, which permits the escape of fluid, closely resembling plasma in composition, into the tissue spaces. The fluid

FIG. 196.—Urticaria Factitia in member of British Merchant Navy. (From a photograph by Dr. B. S. Kent.)

Response to light stroking of skin in sensitive subject (cf. p. 110); note whealing.

contains 4–5% of protein, and clots on standing. As the wheal increases in size it gradually becomes paler owing to compression of the minute vessels. A marked arteriolar flush must be present if effective whealing is to occur, because a rich blood supply is needed to provide the necessary large volume of fluid. The wheal is not due primarily to increased intracapillary pressure, for if the arm is compressed to produce venous congestion the wheals formed are *smaller*, owing to the diminished blood flow, in spite of the fact that the capillary pressure is raised. If trypan blue is injected intravenously into a normal person it does not escape into the tissue spaces, but it does pass into a wheal while it is forming owing to increased capillary permeability.

Triple Response from Histamine.—If *histamine* (even in dilutions of 1 in 3000) is injected into the skin it produces a typical triple response:

(i) Local dilatation of the capillaries and venules by direct action.

(ii) Widespread dilatation of surrounding arterioles from a local axon reflex. The histamine stimulates local afferent nerve endings; impulses pass up the cutaneous afferent nerve (A, Fig. 195) which branch to give collaterals (C, Fig. 195) to supply dilator fibres to neighbouring arterioles. The vasodilators concerned are the same as those which come into action when the peripheral end of a dorsal nerve root is stimulated (p. 308).

(iii) Local increased permeability of the walls of the minute vessels, increased exudation of fluid and wheal formation. If the skin is firmly stroked and simultaneously a histamine solution is inoculated, a similar "triple response" develops in each case; the responses have exactly the same time relations as regards onset and disappearance. The above observations suggest that the triple response elicited by heavy stroking may be due to the *local release in the skin of histamine or some closely related substance* ("*H*"-*substance*).

Evidence for "H"-Substance.—(1) Many agents can produce the triple response: pricking, scratching, freezing, or burning the skin, the passage of strong electric currents, the inoculation of irritants such as 5% HCl, 5% NaOH, 1% morphine, 2% atropine, cantharides, peptone; fish extracts, flea and gnat bites in susceptible subjects; alcoholic extracts of skin, liver, and lung. Many dissimilar agents—*all of which injure the tissue*—thus produce the same complex vascular response; the simplest explanation of this uniformity of action is that they all *release the same chemical substance* in the skin, and it is the latter which sets up the triple response.[1]

(2) If histamine is punctured into the skin, a vivid flare develops as stated, which fades to a considerable extent after 10 minutes. The blood supply to one arm is obstructed, a histamine solution is inoculated, the occlusion of the vessels is kept up for 11 minutes and is then released. The control arm is similarly treated, except that the histamine is inoculated 1 minute before the vessels are released. On release of the circulation the skin of both arms becomes uniformly red (*reactive hyperæmia* (p. 329)), but this redness rapidly fades, except in the regions where the histamine was injected. As soon as the histamine flares are visible their outlines are marked out in ink, and again at suitable intervals as fading takes place (Fig. 197). It is found that the fading in the two arms occurs *simultaneously*. The conclusion is drawn that when histamine is injected into skin in which the circulation is arrested it is *retained* unchanged in the skin for the full period of occlusion and produces its full effects as soon as the circulation is released.

(3) On the basis of these findings a crucial experiment can now be performed. The observations described are repeated, but instead of inoculating histamine, the skin of a sensitive subject is firmly stroked; exactly the same results are obtained as with the histamine injections, *i.e.* the flare develops and fades in the same way in both arms when the circulation is released. The experiment proves that the mechanical stimulus itself does not reflexly produce the flare, because if that were the case, during the 10

[1] It is interesting to note that a *nettle sting* actually contains both histamine and acetylcholine.

minutes of vascular occlusion which followed the application of one of the stimuli, these effects should have passed away. One must conclude that the mechanical stimulus causes the *release of some chemical substance* which behaves exactly like histamine; it is retained intact in the skin during the period of circulatory arrest and produces its effects as soon as the circulation is released.

The evidence so far given may be summarized thus:

(i) A large number of agents which injure the tissues elicit the " triple response."

(ii) The " flare " response to heavy stroking develops in a normal manner after 10 minutes of circulatory arrest, suggesting that a chemical agent acts as an intermediary.

(iii) Responses which are identical in all respects can be obtained by

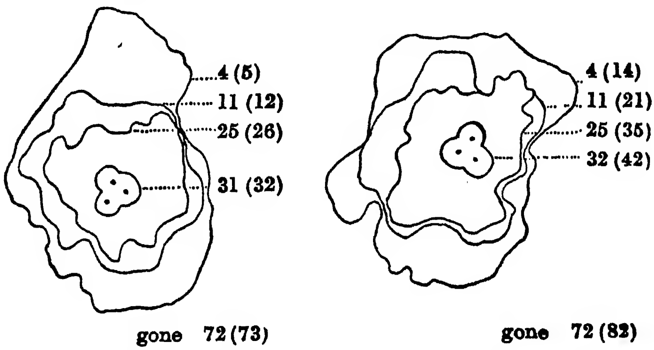


FIG. 197.—Effect of Period of Vascular Occlusion on Flare Response to Injection of Histamine in Normal Subject.

Left outline corresponds to left arm area, right outline to symmetrical right arm area of skin. The vessels of both arms were occluded. A group of three histamine punctures (1/3000 solution) was put down on the right forearm 1 minute after occlusion, on the left forearm 10 minutes after occlusion. The circulatory arrest was continued in each case to the 11th minute, when both arms were released. The outlines of the flares aroused by the stimuli were drawn at the times in minutes after release of vessels noted in figures against the corresponding contours. Numbers in brackets represent corresponding times elapsing after stimuli were laid down; (*e.g.* 4 (14) means outline 4 minutes after release of circulation and 14 minutes after puncture was performed). Position of punctures shown as dots. (Lewis, *Blood Vessels of Human Skin and their Responses.*)

inoculating histamine in high dilution into the skin. There is little doubt therefore that some agent with a histamine-like action (" H "-substance) is released when the skin is injured.

(4) That this " H "-substance is actually histamine itself is suggested by the following observations.

(i) Histamine can be demonstrated in extracts of human epidermis.

(ii) The back of a susceptible subject is extensively and severely stroked till a large wheal develops. Under these circumstances the " H "-substance might be expected to be formed in such amounts that some might get into the circulation and *produce effects at a distance.* It is found that flushing and rise of temperature of the skin of the face occur like those which follow a subcutaneous injection of 0·06 mg. of histamine (p. 337); even stimulation of the secretion of gastric juice may result (p. 338).

The physiology of histamine is fully discussed on pp. 335 *et seq.*

Effect of Heat.—(1) LOCAL EFFECTS.—If the arm is plunged into a bath of water at 40°–45° C., marked reddening of the skin takes place owing to dilatation of the minute vessels. The response is brought about in the following ways:

(i) By the direct action of the rise of temperature on the capillaries and the release of "H"-substance.

(ii) In addition, *reflex* dilatation of the limb arterioles and veins takes place, which leads to an increased blood flow through the skin.

(iii) At still *higher* temperatures a typical flare, wheal, and blister are produced owing to the formation of large amounts of the "H"-substance.

If the vessels of an organ are irrigated with saline at 42° C., the dilatation affects first the capillaries and later the arterioles; similar changes probably occur during tissue activity when the temperature of the active organ rises. As pointed out on p. 18, local heating produces a marked rise of capillary pressure in the skin.

(2) REMOTE EFFECTS.—Heat applied, *e.g.* to the legs, produces dilatation of skin blood vessels elsewhere, especially in the face and hands. The effect is mainly due to the warmed blood acting *directly* on the vasomotor centre inhibiting its "tonic" vasoconstrictor discharge. [A similar effect is produced by introducing warm water into the stomach (where the temperature nerve endings are few).] But as the remote vasodilatation still occurs though to a smaller extent after occluding the blood vessels of the heated limb (to prevent the heated blood from entering the general circulation) a *reflex* effect from temperature nerve endings is also involved.

Fig. 198 shows the effects on the finger circulation produced by plunging both legs into water at 45° C.: there is a rise of blood temperature as shown by the increase in rectal temperature;

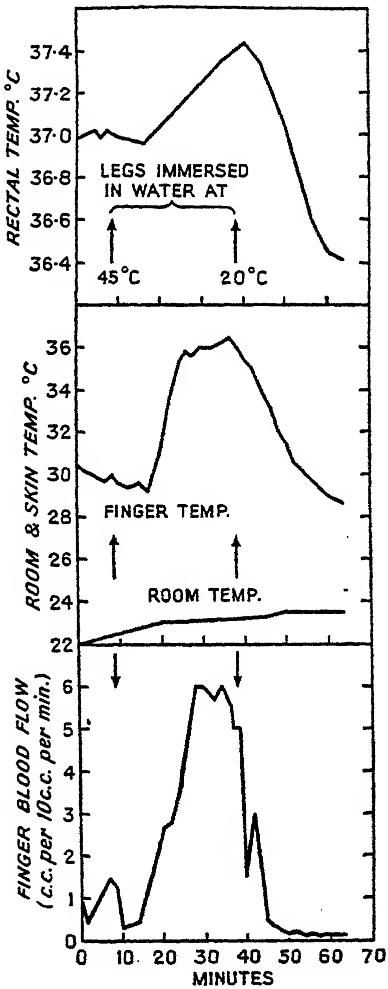


FIG. 198.—Effects of Heating and Cooling Body on Circulation in Fingers. (Wilkins *et al.*, *Clin. Sci.*, 1938, *3*, 407.)

Records from above downwards: rectal temperature (°C.); temperature of skin of finger (°C.) and room temperature (°C.); finger blood flow (c.c./*ten* c.c. of finger volume/minute).

At the first arrow the legs were immersed in water at 45° C. Rectal temperature goes up (from 37° to 37·4° C.). Temperature of tip of right forefinger (Finger Temp.) and blood flow in right middle finger both increase markedly (cf. Fig. 217).

At the second arrow the legs were immersed in water at 20° C. Finger temperature and blood flow and rectal temperature all fall.

the skin arterioles dilate as shown by the increased local temperature and local blood flow. Muscle blood flow, however, is unaltered.

Effect of Cold.—(1) LOCAL EFFECTS.—Cold produces complex effects:

(i) There may be slight pallor of the skin with arteriolar and capillary constriction; this is best seen when skin reddened by warmth is plunged into colder water. *Whitening* of the skin from cold is abnormal, and its production is not fully understood.

(ii) If the hand is immersed into very cold water (5°–10° C.) for 5–10 minutes, the skin becomes a bright *red* colour which is indistinguishable from that produced at a temperature of 45° C. The arterioles are contracted but the minute vessels are dilated. One would have expected cyanosis to develop (cf. (iii)). At temperatures under 10° C., however, oxyhæmoglobin undergoes very little dissociation, and the oxidative activities of the tissues are also very depressed, so that in spite of the inadequate and very slow blood flow through the part, the hæmoglobin is not sufficiently reduced to give a dark tint to the blood.

(iii) At skin temperatures of 20°–25° C. *cyanosis* may develop. The vascular changes are similar to those described in (ii); one must presume that at this critical temperature the blood flow is reduced more than the oxygen utilization by the tissues; consequently the oxyhæmoglobin undergoes extensive reduction and a dusky colour appears.

(iv) *Freezing* the skin results in extensive cellular damage and liberation of " H "-substance with the usual triple response.

(2) REMOTE EFFECTS.—Cold applied to a part (*e.g.* the leg) produces remote effects as well (*e.g.* in the fingers (Fig. 198)). The vasomotor centre is stimulated both directly by the cooled blood reaching it and chiefly reflexly by afferent impulses from the cooled area. The resulting vasoconstriction (especially in the hands and feet) may lead to a rise of blood pressure; (the latter change is measured in the *cold pressor test* [1]); the blood flow through muscles is unchanged.

Action of Light.—Exposure of the skin to strong light produces, as is well known, an erythema. Finsen exposed the skin to the rays from a 40,000 candle-power lamp; after twenty minutes the whole area became uniformly red. This erythema is due to the effect of the *heat*, and diminishes in 2 hours. After 3 hours, erythema due to light appears, reaches its maximum in 12 hours, and decreases after 2 days. The redness is associated with capillary dilatation and stasis. If the skin is previously covered with a *glass* plate, which keeps back the ultra-violet rays, the delayed erythema is not obtained; on the other hand, *rock crystal*, which allows these rays to pass through, affords no protection. Light erythema is thus due to *ultra-violet rays.* Lewis suggests that ultra-violet rays act by causing a slow release of the " H "- substance in the skin. The visible spectrum, however, can also

[1] *Cold pressor test.*—The rise of blood pressure (mainly *reflex* in origin) produced by cold is determined as follows. The subject rests in the supine position for 20–60 minutes in a quiet room so that his blood pressure falls to basal level; a cuff is wrapped round one arm. One hand is dipped into iced water (at 4° C.) to a point just above the wrist; the blood pressure is read at 30 and 60 seconds. The hand is removed from the water and blood pressure readings taken every 2 minutes till the basal level is regained. In normal subjects the maximal rise occurs after 30 seconds and the rise is over in 2 minutes. A rise of more than 20 mm. Hg systolic or 15 mm. Hg diastolic is regarded an excessive reaction and suggests that the subject's vasomotor reflexes are exaggerated.

act on the capillaries if it is employed in sufficient concentration and for a long enough time. The erythema in Finsen's experiment disappeared in 10 days and was followed by pigmentation and desquamation. The effect on the capillaries, however, was of an enduring nature, as shown by the fact that when 6 months later the skin of the arm was rubbed, the areas which had been exposed in the experiment became redder than the other parts of the skin.

Reactive Hyperæmia.—The circulation to a limb is completely obstructed for a period of, say, 5 minutes by means of a suitably inflated blood pressure cuff. The obstruction is then removed. Within a few seconds the skin distal to the block flushes brightly, attaining its maximum colour at once, and becomes hotter; the volume of the pulse is increased owing to lowered local diastolic pressure and consequently greater pulse pressure. The cutaneous hyperæmia declines after 20–30 seconds, and is gone in a few minutes. The blood flow changes can be studied quantitatively by means of venous occlusion plethysmography (p. 305). Fig. 199 shows that after 5 minutes' ischæmia there is an increase in the forearm blood flow from a basal level of

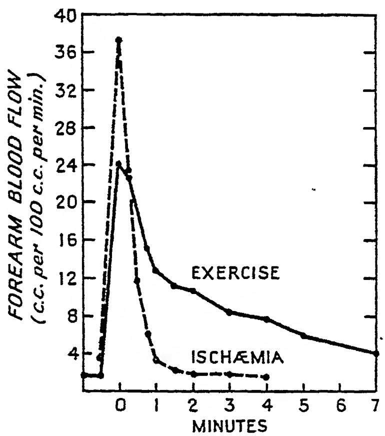


FIG. 199.—Changes in Muscle Blood Flow from Exercise and Ischæmia. Forearm blood flow before and after 4 minutes' continuous exercise (continuous line) and 5 minutes' ischæmia (broken line). Time in minutes after end of ischæmia or exercise. The dilator response to ischæmia is greater but more transient. (Grant, *Clin. Sci.*, 1938, *3*, 170.)

2 c.c. to about 38 c.c. per 100 c.c. per minute. This enormous increase is due mainly to dilatation of *muscle* vessels. The duration of the occlusion determines the degree of resulting hyperæmia; occlusion for 0·5, 1·5, 5, and 15 minutes respectively increased the blood flow in an arm (*per 100 c.c.*) from 1 c.c. to 3, 7, 10, and 12 c.c. per minute (Fig. 200).

During reactive hyperæmia a large increase also occurs in the cutaneous blood flow as can be inferred from the rise of skin temperature; it can be conveniently measured in studies on the hand which consists mainly of skin.

The vessels concerned in reactive hyperæmia are (i) the capillaries which

dilate, accounting for the flushing of the skin, and (ii) the arterioles which dilate, accounting for the increased blood flow and pulse volume and the local rise of temperature. These changes are probably due to the accumulation during occlusion of dilator metabolites, known or unknown. We may go some way towards identifying the agents concerned. If a limb is made ischæmic in animals by blocking the arterial supply, the venous outflow on releasing the circulation is said to contain an excessive amount of members

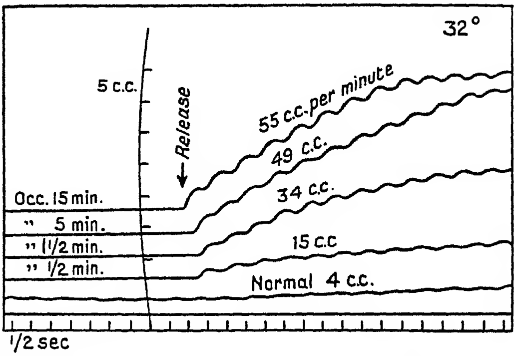


FIG. 200.—Effect of Duration of Occlusion on Arm Blood Flow during Reactive Hyperæmia. (Lewis and Grant, *Heart*, 1925, *12*.)

Arm volume records of normal human subject taken with a plethysmograph. Two cuffs are placed on the upper arm (above the plethysmograph), one above the other; the upper is connected with a reservoir standing at a pressure of 350 mm. Hg and is used to occlude the circulation completely. The pressure in the lower armlet is raised to 70 mm., *i.e.* well above the normal venous pressure. Occlusion is maintained for ½, 1, 5, and 15 minutes respectively, as shown in the records in the Fig. At "Release" (indicated by the arrow) the pressure in the upper armlet is abruptly released. Arterial blood at once begins to flow into the limb, but the venous outflow is prevented by the lower pressure cuff. The subsequent increase in the volume of the limb (rise of the curve) is a measure of the rate of arterial inflow into the limb. It is compared with the lowest curve (normal) in which the upper ("occluding") cuff was not inflated and the pressure was raised only in the lower ("collecting") cuff. The difference in ascent of the control and experimental curves is a measure of the vasodilatation taking place. The *absolute* rates of inflow (in c.c. per minute) were: control 4; experimental 15, 34, 49, and 55. The arm volume was 460 c.c.

of the *adenosine group* and sometimes *histamine*; these, and local acidity, may be the main agents concerned.[1] These substances are possibly normally concerned with regulating the calibre of arterioles and capillaries (pp. 317, 320).

Ischæmia lasting for 10 minutes produces a greater degree of hyperæmia of a limb than does soaking it in water at 43° C. for half an hour. Reactive hyperæmia also follows *venous* obstruction.

FŒTAL CIRCULATION. PATENT DUCTUS ARTERIOSUS

Foetal Circulation.[2]—The peculiarities of the fœtal circulation may be related to two fundamental facts:

(i) The "functioning fœtal lung" is the *placenta*, which is situated on the systemic circulation; the umbilical artery thus corresponds to the

[1] These dilator agents may also be responsible for promoting the establishment of a collateral circulation when the obstruction is of a lasting character.

[2] Barclay *et al.*, *Fœtal Circulation*, Oxford, 1944. Barcroft, *Researches on Prenatal Life*, Oxford, 1946.

postnatal pulmonary artery, and the umbilical vein (which brings oxygenated blood to the fœtus) corresponds to the pulmonary veins.

(ii) It is necessary to supply blood at the highest possible oxygen pressure to the fœtal *brain.* To achieve this end the "arterialized" umbilical vein blood, mixed with the comparatively small venous return from the lower half of the body, passes along the inferior vena cava to the right auricle whence it is mainly diverted via the *patent foramen ovale* [*via sinistra*] into

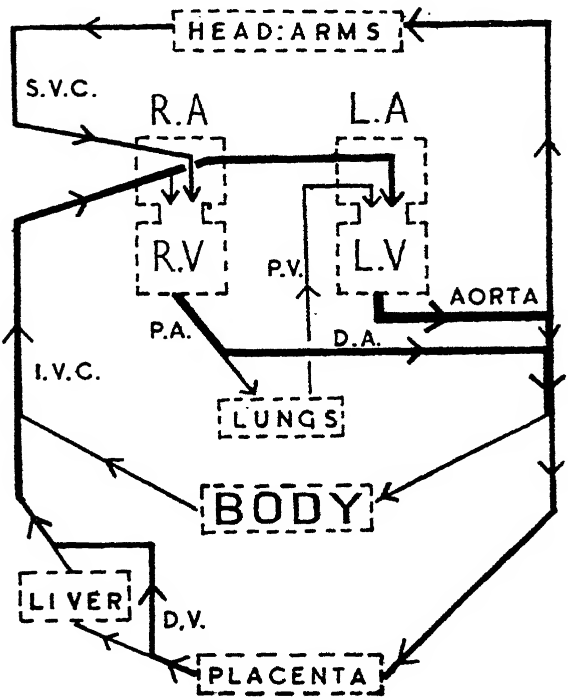


FIG. 201.—Schema of Fœtal Circulation. (Diagram drawn by Prof. David Slome.)

D V.=ductus venosus D.A.=ductus arteriosus; P.A.=pulmonary artery; P.V.=pulmonary vein.

the left auricle. This blood, which is reasonably well oxygenated, passes into the left ventricle, the aorta, and, via its proximal branches, to the *head* and arms.

(iii) As the fœtus, *in utero,* does not breathe, the whole of the output of the right ventricle is not sent to the lungs; most of it is diverted via the *patent ductus arteriosus* to the aorta.

The details of the fœtal circulation are as follows [1] (Fig. 201):

(1) The *oxygenated* blood from the placenta travels in the umbilical vein to the liver. There, part of the blood goes via the *ductus venosus* directly

[1] The account is based mainly on studies of the circulation in the living sheep fœtus but in essentials the description probably applies also to the human fœtus.

to the inferior vena cava; the rest is distributed to the left two-thirds of the liver by offshoots of the umbilical vein; the right one-third of the liver receives the blood from the portal vein. The non-ductus blood leaves the liver in the hepatic veins. The inferior vena cava thus receives:

(i) oxygenated blood via the ductus venosus,
(ii) less oxygenated blood via the hepatic veins,
(iii) very reduced blood from the hind part of the body.

(2) In the *right* auricle, the inferior caval blood is divided into two streams, a large left and a small right one.

(i) The large left stream which consists mainly of oxygenated ductus venosus blood, passes through the foramen ovale [via sinistra] into the *left* auricle; there it is joined by a small volume of reduced pulmonary venous blood. The left auricle discharges into the left ventricle which expels its (oxygenated) blood into the aorta, whence it passes mainly to the coronaries and via the carotids to the *brain*; it also goes via the subclavians to the arms.

(ii) The small right stream is joined by the very reduced blood from the superior vena cava and the coronary sinus. It is discharged from the right auricle into the right ventricle and thence into the pulmonary arteries. Most of the blood is diverted via the ductus arteriosus into the aorta; the smaller residue flows through the lungs. At their point of union the ductus arteriosus and aortic arch are approximately of equal diameter.

(3) Beyond its junction with the ductus arteriosus the aorta contains a mixture of blood, most of it reduced, which has entered via the ductus arteriosus; the balance is oxygenated blood which has flown on from the more proximal part of the aortic arch. This mixed aortic blood which is poorly oxygenated is supplied to the distal parts of the body (abdominal viscera, body wall and hind limbs).

CIRCULATORY CHANGES AT BIRTH.—During, or shortly after, birth important circulatory changes take place.

(1) *Closure of the Umbilical Vessels.*—This occurs as a result of active contraction of the smooth muscle coat of these vessels; the nature of the stimulus is not known, but it is not nervous as the umbilical vessels are devoid of a nerve supply. If the vessels are torn across the mechanical stimulus promotes a powerful contraction of the vessel walls. During the time that elapses between birth and the closure of the umbilical vessels, the greater part of the blood in the cord and placenta is transferred to the child. The cord should not be tied until pulsation in it has ceased so as not to deprive the infant of this extra supply of blood. With the cessation of the venous return along the umbilical vein the volume of the liver and spleen falls; the portal vein blood now supplies the whole liver instead of one third only. The smooth muscle sphincter at the junction of the ductus venosus and the umbilical vein contracts and prevents any backflow of venous blood into the umbilical vein.

(2) The *foramen ovale* [via sinistra] is closed owing to the folding up and apposition of its valve. This closure follows the onset of respiration and always precedes the closure of the ductus arteriosus.

(3) The *ductus arteriosus* closes by contraction of its thick muscle coat a few minutes after breathing has commenced. This response is the result of oxygenation of the blood and not of gaseous distension of the lungs.

Closure does not occur if the lungs are inflated with nitrogen, but it does occur if oxygen is injected intravenously; nerve stimulation (of vagus or sympathetic) has no effect on the ductus. Later the vessel becomes fibrosed and permanently obliterated; should normal closure fail to occur (for unknown reasons) the condition of *patent ductus arteriosus* results.

Circulation in Cases of Patent Ductus Arteriosus.[1]—(i) As after birth the aortic blood pressure far exceeds the pulmonary, the direction of flow in the patent ductus arteriosus is from the aorta to the pulmonary artery (*i.e.* the reverse of that in the fœtus). A proportion (often more than half) of the left ventricular output is shunted into the pulmonary circulation. The effects on the systemic circulation resemble closely the findings in a case of gross arterio-venous aneurysm, and to a considerable extent those in aortic regurgitation (p. 298).

(ii) The following data illustrate quantitatively what may occur: the venous return from the periphery ("peripheral blood flow") is 6 litres per

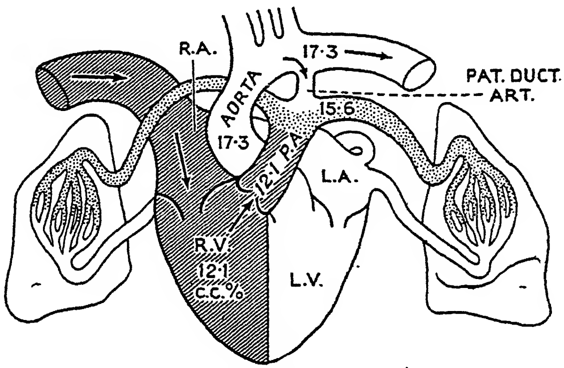


FIG. 202.—Circulation in Patent Ductus Arteriosus. (After Eppinger, Burwell, and Gross, *J. clin. Investig.*, 1941, *20*, 128.)

minute; this volume of mixed venous blood is received by the right auricle and pumped out by the right ventricle into the pulmonary artery. Through the ductus the pulmonary artery receives say 12 litres per minute of almost fully oxygenated arterial blood from the aorta. The total blood flow through the lungs is thus 18 litres per minute; the aortic "shunt" blood undergoes no further change in the lungs, where the mixed venous blood (from the right ventricle) is oxygenated in the usual way. Judging from animal experiments in which a comparable condition can be established by anastomosing the subclavian to the pulmonary artery, the blood pressure in the latter rises surprisingly little, *e.g.* from 16 to 27 (later falling to 22) mm. Hg, presumably owing to the readily distensible character of the pulmonary vessels; but some rise of pulmonary pressure does occur which increases the work of the right ventricle. The "venous" return to the left auricle equals the "total pulmonary blood flow," *i.e.* 18 litres, which is also the output of the left ventricle.

[1] Gross, *J. Amer. med. Assoc.*, 1940, *115*, 1257; Eppinger and Burwell, *ibid.*, 1262.

(iii) This enormous increase in the work of the left ventricle (equivalent to that occurring in free aortic regurgitation) leads to an increase in its size; the left auricle enlarges for the same reason. Of the 18 litres entering the aorta, 12 litres (as already mentioned above) are shunted via the patent ductus into the pulmonary artery and only 6 litres pass to the peripheral circulation to supply the tissues.

(iv) The leak into the pulmonary artery (from the aorta) which continues throughout the cardiac cycle gives rise to a continuous hum and thrill in the second interspace, to prominence and increased excursion of the pulmonary artery, and to congestion of the pulmonary bed.

(v) The systolic blood pressure rises little in spite of the great left

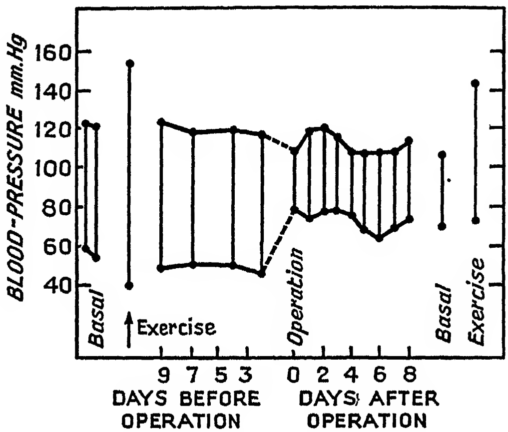


FIG. 203.—Systolic and Diastolic Blood Pressure at Rest and after Exercise before and after closing ductus in case of Patent Ductus Arteriosus. (After Eppinger *et al.*, *J. clin. Investig.*, 1941, *20*, 141.)

Note that after tying the ductus the diastolic and pulse pressures fall.

ventricular output because the leak out of the aorta occurs throughout systole; diastolic pressure, however, does fall (*e.g.* to 40–50 mm. Hg) owing to the abnormal leak during diastole. The other arterial phenomena of arterio-venous aneurysm (detailed on p. 298) may be present.

(vi) The O_2 content of the blood in the ventricles, aorta, and pulmonary artery, proximal and distal to the point of entry of the ductus arteriosus, in a typical case of patent ductus is shown in Fig. 202. The blood in the right ventricle and the proximal part of the pulmonary artery had an O_2 content of 12·1 c.c.%. The blood in the left ventricle and aorta had an O_2 content of 17·3%. Owing to the large volume of arterial blood which flowed from the aorta along the ductus into the pulmonary artery, the O_2 content of the blood in the distal part of the pulmonary artery rose to 15·6%.

(vii) After the patent ductus arteriosus has been ligatured at operation, the thrill and murmur disappear, the heart becomes smaller, the diastolic blood

pressure rises (Fig. 203) and the left ventricular output falls to equal that of the right ventricle.[1]

PHYSIOLOGY OF HISTAMINE. TRAUMATIC SHOCK

Histamine.[2]—Many tissues contain histamine; species variations are marked but usually the concentration is highest in the lungs and skin. This histamine is normally combined in the tissues in some form which prevents it acting locally or being released into the circulation. When the tissue is acted on by appropriate stimuli, some of its bound histamine content is released in a pharmacologically active form. Blood histamine is chiefly present in the cells (granulocytes and platelets); the amount in the plasma is minute.[3]

Histamine is β-iminazolyl-ethylamine; the body store of histamine is derived originally from the important amino-acid histidine (iminazolyl-α-amino-propionic acid) by removal of CO_2. This decarboxylation may result from the action of bacteria on the food before it is eaten, or may occur in the bowel; the histamine thus formed is absorbed chiefly from the

[1] The "peripheral blood flow," "total pulmonary blood flow," and the "shunt" as defined above can be determined at operation in man by collecting (by direct puncture) the following blood samples and determining their O_2 content:

(*a*) Mixed venous blood from the right ventricle, 12·1 c.c.% (=V).

(*b*) Arterial blood from the femoral artery, 17·3 c.c.% (=A).

(*c*) Blood from the pulmonary artery distal to the patent ductus (mixture of mixed venous (right ventricle) and arterial (aortic) blood), 15·6 c.c.% (=M).

The O_2 consumption per minute is determined in the normal way (p. 375), and is, *e.g.*, 312 c.c.

(1) Right ventricular output (=" peripheral blood flow") is determined by the usual Fick principle formula (p. 278), *i.e.*:

$$\frac{O_2 \text{ consumption}}{A-V}\times 100=\frac{312}{17{\cdot}3-12{\cdot}1}\times 100=6 \text{ litres per minute.}$$

(2) To determine the degree of "shunt" and consequently the left ventricular output the following facts must be considered. In the pulmonary artery each 100 c.c. of right ventricular blood (*i.e.* V) containing 12·1 c.c.% of O_2 mixes with an unknown volume (x c.c.) of aortic arterial blood containing 17·3 c.c.% of O_2 (=A) to give $100+x$ c.c. of total pulmonary blood flow containing 15·6 c.c.% of O_2 (=M). The following formula (for which I am indebted to Sandra Wright) can be used to determine x:

$$A\left(\frac{x}{100+x}\right)+V\left(\frac{100}{100+x}\right)=M.$$

$$\therefore\ x=\frac{100\,(M-V)}{A-M}.$$

Substituting the values quoted,

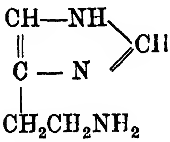
$$x=\frac{100\,(15{\cdot}6-12{\cdot}1)}{17{\cdot}3-15{\cdot}6}=205 \text{ c.c. (say 200 c.c.).}$$

In this example the volume of shunt via the patent ductus (in relation to every 100 c.c. of right ventricular outflow) is 200 c.c., *i.e.* twice the simultaneous right ventricular output. As the latter is 6 litres per minute, the shunt flow is 12 litres per minute and the total pulmonary blood flow is 18 litres per minute, which equals the left ventricular output.

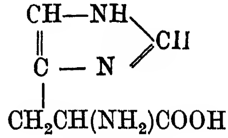
[2] Best and McHenry, *Physiol. Rev.*, 1931, *11*, 371. Symposium, *Ann. N.Y. Acad. Sci.*, 1950, *50*, 1013–1208.

[3] Histamine can also be extracted from certain peripheral nerves which are believed to be *histaminergic*, *i.e.* produce peripheral effects by releasing histamine at their endings.

ileum. A specific histamine decarboxylase is present in some organs and it may form histamine from histidine locally.



Histamine



Histidine

Histamine is rapidly destroyed by an oxidase called *histaminase* which is present in many tissues and especially in the kidney and placenta. Histaminase is, however, not specific in its action; it also acts on many

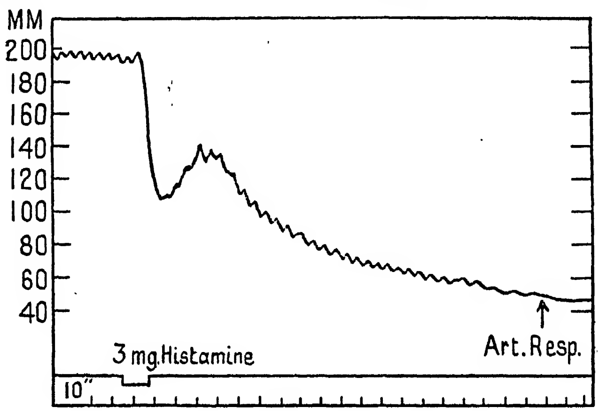


FIG. 204.—Action of Histamine on Circulation.

Cat. Record of blood pressure. Inject 3 mg. of histamine. The second fall lowers the pressure from 140 to 50 mm. and leads to respiratory failure. (Dale and Laidlaw, *J. Physiol.*, 1919, *52*.)

other diamines, *i.e.* it is really a diamino-oxidase. The blood histaminase is greatly raised in pregnancy, its source being the placenta.

Histamine is related to various bodily processes: (i) it may be a normal *dilator* metabolite which is released during *tissue activity* or during *ischæmia*; (ii) it is released when certain tissues are *injured*, in response to *infections* and in *anaphylaxis*. It is convenient to consider its pharmacological actions in detail first; the species differences are many and important.

(1) CIRCULATION.—Intravenous injection of histamine in minute doses produces a *marked fall of blood pressure* [1] (Fig. 204). The classical studies were carried out in the cat in which species it can be shown that histamine dilates the capillaries but constricts the arterioles. The circulatory effects in the cat are as follows:

(i) *Initial Pulmonary Constriction.*—The sharp *initial* fall of pressure is due to temporary constriction of the pulmonary arterioles resulting in interference with the output of the right ventricle; there is consequently a diminished inflow into the left ventricle and a smaller output into the

[1] In the rabbit a *rise* of blood pressure occurs.

systemic circulation. The right ventricle soon recovers and resumes its unembarrassed activity, and the blood pressure may recover slightly.

(ii) *Phase of Capillary Paralysis.*—There is secondarily a steady prolonged fall of blood pressure which may lead to circulatory failure, inadequate blood flow to the brain, and cessation of respiration. This *second, lasting, depressor effect* is mainly or wholly due to *capillary paralysis* and occurs in spite of *arteriolar constriction.*

The evidence (in the cat) is as follows:

(*a*) *Arteriolar constriction.*—A purely arterial system can be prepared by cutting the mesentery at its attachment to the intestine; the superior mesenteric artery and its branches are then perfused, and it is found that histamine diminishes the rate of escape of fluid from the cut end, indicating that it constricts the arterioles.

(*b*) *Heart.*—Histamine does not directly depress the activity of the isolated heart perfused through the coronary vessels; if the heart is examined in the intact animal poisoned with histamine, it is found to be beating forcibly. In spite of this, however, the *cardiac output* (in the second stage) *is very small.*

(*c*) *Capillary paralysis.*—The fall of blood pressure has thus occurred in spite of arteriolar constriction and a forcibly acting heart; the obvious cause of the fall of blood pressure is the *poor cardiac output.* As this is not due to a primary cardiac disorder it must be the result of a failure of the venous return which in turn results from generalized capillary paralysis, which can be readily demonstrated as follows:

(α) In spite of the arteriolar constriction the intestines, pancreas, and other solid viscera are *intensely congested* and œdematous.

(β) The nerves going to the limb of a cat are severed; the arterioles relax (because the vasoconstrictor nerves have been cut) and the limb becomes warm owing to the larger blood flow through it. The pad of the foot remains pale, indicating that the capillaries are contracted in spite of full arteriolar dilatation (cf. p. 321). The injection of histamine at once makes the pad flush because it paralyses the capillary wall, which now gives way before the arteriolar pressure.

(γ) If a histamine solution is dabbed on to an organ like the pancreas it becomes flushed with blood; as the arterioles are contracted, the increased vascular content of the organ must be due to capillary dilatation.[1]

(δ) *Decreased plasma volume.*—Histamine increases the *permeability* of the capillaries (p. 13) so that protein and fluid escape from the vessels into the tissue spaces, still further reducing the volume of circulating plasma (and further aggravating the circulatory collapse). The red blood count per c.mm. and the hæmoglobin percentage consequently rise.

Action in Man.—In man (and also in the dog and monkey) histamine produces *arteriolar dilatation in addition to capillary paralysis.* A subcutaneous injection in man of 0·3 mg. of histamine causes general flushing of the skin, a rise in its temperature of 1° or 2° C., a small decline of systolic, a great fall of diastolic, pressure, and a rise of pulse rate; the flush is most conspicuous in those parts that are previously red. Even doses of 0·06 mg.

[1] The coronary vessels may be constricted or dilated (depending on the species); both pulmonary arteries and veins are constricted (as noted on p. 336). In the dog, histamine contracts the ring of muscle at the point of entrance of the hepatic vein into the inferior vena cava with resulting intense *engorgement of the liver* and other abdominal viscera.

cause flushing of the skin of the face and a rise of 0·5° C. in its temperature (cf. p. 326).

There is dilatation of the vessels of the brain and a rise of cerebrospinal fluid pressure. This leads to stretching of the dura mater and consequent stimulation of the local sensory endings of the fifth nerve. These changes may account for the *headache* which so regularly follows injection of histamine in man.[1]

The action of histamine on the human skin is discussed on p. 325.

(2) SMOOTH MUSCLE.—This is contracted throughout the body, including the muscle coat of the bronchioles, intestine, spleen, and uterus. In the guinea-pig the bronchial spasm may cause death from asphyxia (Fig. 205).

(3) GLANDS.—There is increased secretion of saliva and a very marked flow of *gastric juice* (the latter response forms the basis of the "histamine test" (p. 783)). Repeated injections may produce gastric ulceration; if the gastric mucosa is damaged, histamine may delay healing. It has no *direct* secretory effect on the pancreas; increased pancreatic secretion is produced by histamine only when the pylorus is open but not if it is tied off.

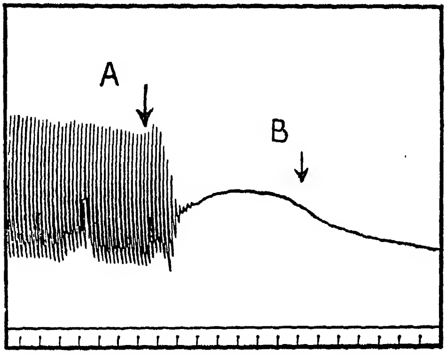


FIG. 205.—Action of Histamine on Bronchial Calibre.

Pithed guinea-pig. The expansion of the artificially inflated lungs is recorded graphically and varies directly with bronchiolar calibre. At A inject 0·5 mg. histamine. The bronchioles close completely as shown by the cessation of rhythmic expansion of the lungs. At B inject 5 mg. atropine, which is quite ineffective in overcoming the spasm. (Dale and Laidlaw, *J. Physiol.*, 1910.)

Rôle of Histamine in the Body.—(1) RELATION TO INJURY.—It is released when human skin is injured and is responsible for the vascular changes constituting the *triple response* (p. 325). Its doubtful rôle in *traumatic shock* is considered on p. 343; in *anaphylactic shock* and *bacterial infections* (*infra*); in *burns* on p. 344.

(2) DILATOR METABOLITE.—It is claimed that histamine is released from skeletal muscle during activity; it may be an important metabolite formed generally during normal tissue activity. It may also be one of the dilator agents responsible for the vascular changes of *reactive hyperæmia* (p. 330).

(3) RÔLE OF HISTAMINE IN ANAPHYLACTIC SHOCK[2] AND INFECTION.—Anaphylactic shock is produced as follows: a small amount of a foreign protein (the "antigen," or "antigenic protein") is injected into an animal; some days or weeks afterwards the *same* protein is injected a second time. The animal becomes gravely ill and may die even when minute doses of the protein are used on the second occasion, doses that are perfectly harmless when given as a first injection. The first injection of "antigen" thus produces a state of *hypersensitivity* or *anaphylaxis*, which is demonstrated by the appearance of serious symptoms of anaphylactic shock when the injection is

[1] Hess and Pickering, *Clin. Sci.*, 1933, *1*, 77.
[2] Dragstedt, *Physiol. Rev.*, 1941, *21*, 563.

repeated. The response is highly specific and exactly the same protein must be used on the two occasions.

The chief symptoms of anaphylactic shock differ markedly in different species; on the other hand, the symptoms of anaphylactic shock *resemble those of histamine poisoning in the same species.* Thus in both conditions intense bronchoconstriction develops in the guinea-pig; in the dog both produce constriction of the hepatic veins leading to engorgement of the liver; in the rabbit both produce constriction of the pulmonary arterioles. Anaphylactic shock is accompanied by the release of histamine: (i) the injection of the specific antigenic protein into a sensitized animal leads to the appearance of histamine in the blood. (ii) If the antigenic protein is added to the fluid which is perfused through the lungs of a previously sensitized animal, histamine is released and appears in the outflow fluid. This histamine is not newly synthesized but is released during the experimen from some closely related precursor.

The results described may be interpreted as follows: the first injection of the antigenic protein leads to the formation of a specific antibody which is found in the *cells* of different organs. When the antigen is injected a second time, it reaches the tissues and there reacts with the antibody; following this union toxic substances are liberated but their exact source is unknown. These toxins damage the cells in which they appear; one of the results of such tissue injury here as elsewhere (*e.g.* in the skin) is the release of histamine; other substances are doubtless also formed. If the histamine is released in the lungs, it produces characteristic bronchoconstriction and swelling of the lungs owing to exudation of fluid into the alveoli from the poisoned and more permeable capillaries.

On this theory the symptoms of anaphylactic shock result from: (i) *widespread tissue damage* resulting from the action of the locally released toxic product of the antigen-antibody reaction; (ii) *histamine poisoning,* both in the organs in which it is primarily released and elsewhere as a result of its absorption into the general circulation.

When certain *bacterial toxins* (*e.g.* staphylococcal), or snake *venoms* [1] which damage tissues are perfused through the lungs, they may cause histamine to be liberated. But again it must be stressed that the toxins and venoms also damage the tissues directly and in ways not involving the mediation of histamine.

Traumatic Shock.—When a clinician describes a patient as being in a state of "shock," he is not diagnosing a specific disease but is employing a convenient label to attach to a clinical syndrome, *i.e.* to a fairly clearly defined group of symptoms and signs. "Spinal shock" is of course quite a distinct entity with characteristic features described on p. 690. The condition of "shock" is a serious one involving danger to life. It is unfortunately true to say that the diagnosis of "shock" is so loosely and variously used that it may fail to give an outsider a clear indication of how ill the patient is, what symptoms and signs he displays to the bedside observer or what treatment he requires. The clinical state termed "shock" may be associated with a wide variety of causal agencies: among the more important are trauma, hæmorrhage, severe dehydration (*e.g.* from grave diarrhœa or incessant vomiting),

[1] Feldberg and Kellaway, *Austr. J. exp. Biol. med. Sci.*, 1937, *15*, 159, 441, 461; *J Physiol.*, 1937, *90*, 257; 1938, *94*, 187.

severe burns, toxæmias, high-voltage electrical currents, visceral perforation, coronary thrombosis, introduction of foreign proteins into the body, pleural puncture, and great mental stress. It is proposed to discuss here mainly shock due to trauma. It is customary to divide traumatic shock into *primary* (or *immediate*) which comes on at once, and *secondary* (or *delayed*) shock which either follows on the primary or develops after a latent period without manifest signs.

Clinical Picture of Traumatic Shock.[1]—Following a *severe* injury the patient shows " a rapid thready pulse, a pallid or greying or slightly cyanotic appearance of the skin which is cold or moist, sweating, marked thirst, superficial rapid respiration, commonly vomiting, restlessness, lessened sensibility, a dulled mental state," and a subnormal temperature. *Milder* cases of shock do not show all these symptoms, and serious cases may vary in some of their clinical manifestations. Thus the typical picture of the wounded after the Battle of Alamein was that of " a talkative, even garrulous, man, with ashen-grey face and beads of sweat on the brow, tiny pupils narrowed by morphine, making restless, fidgety movements, keeping an apprehensive eye on his bearers lest his wound should be jarred, asking constantly for drinks, and vomiting without warning a few minutes after each drink."

(1) Blood Pressure.—A *fall of blood pressure* is commonly noted and is widely regarded as an indispensable feature of " shock." Some would say " no low blood pressure, no shock," but this attitude is not in accordance with the evidence. Clinically obvious and severe " shock " as defined above may be associated with a blood pressure within the normal range, though far more commonly the blood pressure is reduced, even to levels that make measurements impossible. It will be shown later that after an injury there are a number of factors in operation that *tend* to lower the blood pressure; but the blood pressure only *actually* falls when the elaborate compensatory mechanisms of the body have been overcome. When clinical shock is associated with a normal blood pressure there is usually intense vaso-constriction in certain parts of the body, with resulting decrease in local blood flow and anoxia. When the blood pressure in a case of shock is low there is probably a more generalized inadequate blood supply to the organs. It should be remembered that the organs that are least fitted to withstand severe anoxia of any duration are the brain, the heart, and the kidneys. Up to a point the effects of anoxia on the organs are reversible and can be fully recovered from if the circulation is promptly restored; but if the anoxic state persists for too long a period, irreversible changes occur and recovery will not take place in spite of treatment.

(2) Heart Rate.—Though the *pulse* is commonly rapid it may be slow or within normal range; thus five cases of shock with a blood pressure of 55 mm. Hg or less, had pulse rates of 86, 82, 92, 104, and 76 per minute.

(3) Blood.—The *blood volume* is commonly reduced. When this is due mainly to hæmorrhage, there is a secondary flow of tissue fluid into the circulation, producing a fall of hæmoglobin concentration and of hæmatocrit value (*hæmodilution*). When the decrease in blood volume is due to plasma or simple fluid loss (rather than loss of whole blood) there is *hæmoconcentration*.

(4) Kidneys.—In shock due to crush injuries (p. 343) gross anatomical changes occur in the kidneys. But in other groups of traumatic shock it

[1] Cope, *Lancet*, 1944, i, 702.

has been found that the renal blood flow, the volume of glomerular filtrate and volume of urine excreted are reduced roughly proportionally to the degree of shock. Cortical ischæmia may occur from closure of the interlobular arteries; the blood is then shunted via the juxta-medullary glomeruli into the renal medulla (p. 26).

Mechanism of Shock.[1]—Traumatic shock is thus a general bodily disturbance following an injury which may be limited in extent. The local injury might produce the widespread changes noted in one or both of two ways:

(1) NERVOUS: *i.e.* afferent impulses arising in the injured area, reflexly via the central nervous system, modify in a harmful manner the activities of various organs or systems.

(2) HUMORAL: (i) by (*a*) local loss of blood, plasma, or fluid, leading to decreased circulating blood volume; (*b*) local loss of heat leading to subnormal body temperature; (*c*) local loss of some essential blood constituent.

(ii) By absorption of toxic agents from the injured area into the general circulation.

In *immediate* shock, it is likely that nervous reflexes play a predominant part producing, it is believed, reflex vasodilatation, often cardiac slowing and reduced cardiac output and consequently reduced blood flow and thus tissue anoxia. Unfortunately few detailed analyses of cases of this type of shock have been carried out; but if the suggestions just made are substantiated, immediate shock would turn out to be closely related to the vaso-vagal syndrome (p. 271) or fainting attack associated with slow heart and vasodilatation. In *delayed* shock some or all of the factors enumerated above are probably in action to a varying extent in different cases (*v. infra*). On ultimate analysis the state of shock is due to depressed or disordered tissue metabolism; this in turn may be due to anoxia, toxic agents, or lowered temperature. The anoxia in its turn is due to circulatory failure which might be due to diminished circulating blood volume, to impairment of the heart's action, or to decrease in the peripheral resistance.

Animal experiments planned to analyse the causes and manifestations of shock suffer from obvious, serious, and not wholly avoidable defects. There is first the species difference, as animals may differ in their responses from man. The experiments are usually performed under anæsthesia, which eliminates the factor of conscious pain from the clinical picture, though it introduces a complicating factor of its own, as prolonged and deep anæsthesia can ultimately produce circulatory failure. The experimentalist further does not use the same criteria of "shock" as does the clinician; he does not judge the entire clinical condition, but regards a lowered blood pressure as an indispensable manifestation and in fact as the main index of the severity of the state of shock. The type of injury employed to produce shock has varied very widely and the results in different series of experiments have been regarded as comparable though the procedures have been significantly different. These remarks may help to explain the present state of confusion of the subject and the difficulty of correlating experimental and clinical findings.

Rôle of Specific Factors in Clinical Shock.—(1) HÆMORRHAGE.—The clinical results of severe hæmorrhage are indistinguishable from those of shock; *i.e.* hæmorrhage alone may be responsible for a state of shock. In

[1] Grant and Reeve, *General Effects of Injury in Man*, H.M. Stationery Office, 1951.

fact most cases of clinical shock show changes indicating that considerable blood loss has occurred. In one group the outstanding features were a decrease in the blood volume, a lowered hæmatocrit value (indicative of hæmorrhage with partial compensation by tissue fluid inflow), lowered venous pressure, decreased cardiac output, low blood pressure, and renal insufficiency. In a representative case a few hours after the injury the cardiac output was reduced from (the theoretically expected) 5·3 litres to 3·75 litres per minute; the blood volume was down from 4,500 to 3,500 c.c. and the hæmatocrit value from 45% to 28%. The mean arterial blood pressure (*i.e.* mean of systolic and diastolic pressures) ranged from 30 mm. Hg 3 hours after the injury to 80 mm. Hg 2 hours later. The rate of glomerular filtration was lowered from 110 c.c. to 20–30 c.c. per minute and the renal plasma flow from 600 c.c. to 150 c.c. per minute. The calculated resistance to the blood flow in the kidney was increased fourfold, *i.e.* there was intense renal vaso-constriction. The flow of urine was only 3 c.c. (normal 50 c.c.) per hour. The *total* peripheral resistance was not significantly altered. Calculation shows that the total red cell volume before the injury was 2000 c.c.; 3 hours after, it was down to 1000 c.c. Thus 1000 c.c. of red cells had been lost as a result of the injury, doubtless together with over 1000 c.c. of plasma; in other words it appears that this patient had suffered a hæmorrhage of over 2000 c.c. in a short period; as the total blood volume had only fallen by 1000 c.c., about 1000 c.c. of tissue fluid had been drawn back into the circulation. It is not unlikely that this very large hæmorrhage was responsible for all the other manifestations noted.

In our studies of hæmorrhage in man, we dealt with the results of loss of blood of about 1000 c.c. (p. 81); nothing is known about the effects in man of an acute hæmorrhage of twice that size. In the 1000 c.c. hæmorrhage the peripheral resistance is initially greatly increased because of compensatory vasoconstriction; when fainting occurs it is associated with a collapse of the blood pressure and a great decline of the peripheral resistance from its initial high value. In these cases of shock no increase of total peripheral resistance was noted; if it was present initially it had soon passed away. The low cardiac output, coupled with a normal peripheral resistance, produces the characteristically low blood pressure. The circulatory collapse must be attributed to a primary failure of the venous return which leads to a decreased cardiac output. The only region where marked vasoconstriction was present during shock was the kidney; the very low blood flow through the kidney (plus a possible cortical ischæmia) accounts for the negligible urinary output. More direct evidence is still needed to prove that in cases of the type described hæmorrhage of the extent calculated has really taken place. Once the state of shock develops the vital organs suffer progressively; the decreased cerebral flow depresses first the higher centres and later the vital medullary centres; the decreased coronary flow impairs the contractility of the heart; the decreased renal flow finally produces the manifestations of latent uræmia (p. 73).

(2) Plasma or Fluid Loss.—A large loss of plasma alone or of fluid from the blood could produce a circulatory picture very similar to that described above. It has been suggested that in many injuries, apart from the frank hæmorrhage from torn vessels, there may also be leakage of protein-rich fluid out of vessels whose permeability has been increased. One can only

be certain that plasma loss is the dominant factor when the blood shows hæmoconcentration; this latter finding is uncommon following trauma. Plasma loss on a large scale does, however, occur from a burnt area. Severe vomiting or sweating would still further decrease the plasma volume (producing anhydræmia). There is no evidence that capillaries remote from the traumatized area show any increase in their permeability.

(3) CRUSH SYNDROME.—The victims of air-raids are often buried under heavy masses of masonry which may pin down and crush a limb for a number of hours. When rescued, the patient's clinical state may not be alarming, but progressive deterioration sets in and some two-thirds of the cases die in spite of treatment. Owing to the prolonged complete ischæmia, extensive necrosis develops in the affected muscles and skin. The locally injured vessels permit the leakage of blood or plasma with resulting decreased blood volume, and the limb becomes enormously swollen and tense with impairment of its circulation. The activity of the kidneys is depressed: (i) by the decreased blood flow which is due partly to the lowered blood pressure and partly to compensatory vasoconstriction; (ii) there may be cortical ischæmia which cuts the glomeruli out of the circulation; (iii) in addition *myohæmoglobin* is absorbed from the damaged muscles and while being excreted by the kidneys is commonly precipitated in the renal tubules, blocking the escape of urine; the renal tubules often show extensive degenerative changes. The volume of the urine is decreased greatly, and myohæmoglobin may be present in it. Uræmic symptoms from renal failure complicate the clinical picture.[1]

It is not improbable in view of the crushed state of the muscles that various products of tissue damage are absorbed from the limb and produce systemic toxic effects; but in spite of intensive search, no such substances have yet been demonstrated in the blood. The agents looked for have been mainly substances of the histamine class (which produce capillary paralysis) or general vasodilators like adenosine and related compounds.

(4) REFLEX FACTOR.[2]—In some cases, typically following fracture of a long bone like a femur, shock may be severe without evidence of great damage to soft parts or fluid or blood loss. Experimental analysis suggests that the manifestations of shock in these cases are reflexly produced and may be prevented or greatly reduced by denervation of the injured part or blocking the reflex arc in the spinal cord by spinal anæsthesia. It is not yet known how the afferent impulses set up in the injured part cause the permanent depression of the circulation. In conscious subjects it is well known that severe prolonged pain from any cause can ultimately induce a state of circulatory collapse and marked depression of general nervous activity. (Cf. vaso-vagal attacks, p. 271.)

(5) FALL OF TEMPERATURE.—Temperature regulation becomes defective; exposure to cold (a complication commonly occuring with wounded men) and marked sweating will intensify heat loss and set up a vicious circle in which the body temperature progressively sinks.

(6) INFECTION.—If obvious infection of the injured part supervenes, toxic products formed by the bacteria or released from the damaged tissues may be absorbed and poison the heart and paralyse the peripheral blood vessels leading to circulatory collapse. It is claimed that experimentally in

[1] Bywaters and Dible, *J. Path. Bact.*, 1942, *54*, 111.
[2] Slome and O'Shaughnessy, *Brit. J. Surg.*, 1936, *22*, 589; 1938, *25*, 900.

shock due to crushed limb or hæmorrhage, organisms enter the blood stream and blood cultures are positive for the enteric group, *B. ærogenes* and *B. coli*.

(7) Other *contributory factors* which sometimes play a part are *over-ventilation* leading to CO_2 lack and its sequelæ (p. 406), *emboli* of fat liberated mainly from the marrow of damaged bones, or a rise of plasma *potassium*. In shock increased utilization of *adrenal corticoids* occurs, leading to exhaustion of the gland.

Evidence from the Results of Therapy.—It is not proposed to consider the detailed management of a clinical case of shock; but the results of treatment (when carefully controlled) may throw light on the nature of the disturbances involved. In many clinical cases measures directed to increasing the volume of the circulating blood have proved highly beneficial. These consist of intravenous transfusion of whole blood, plasma, or serum which are retained in the circulation (because of their protein osmotic pressure) far better than saline (Fig. 48). If bleeding is not fully controlled or if, because of altered capillary permeability, protein-rich fluid continues to escape into the tissue spaces, very large transfusions may have to be given. These results support the view that a decrease in the *total* blood volume, or in the *circulating* blood volume, or a disproportion between the capacity of the vascular bed and the total blood volume are often important factors in promoting the circulatory failure that is such a common and considerable factor in shock. Sedatives are also widely employed and prove beneficial probably because they relieve pain and depress certain reflex arcs which were operating harmfully. Adrenal cortex extracts have been given to relieve the supposed state of cortical deficiency (*supra*); but the doses used have probably been too small and there is inadequate evidence that they have proved beneficial clinically; cortisone may perhaps prove useful. The subnormal body temperature in shock has been regarded as an indication for warming the patient; it is, however, dangerous to overheat the patient, as the rise of temperature, by producing vasodilatation may annul the beneficial effect of the pre-existing compensatory vasoconstriction. Heating also increases sweating, thus increasing fluid loss and further diminishing the plasma volume. It must be borne in mind that a low *skin* temperature does not necessarily indicate that the *internal* temperature has fallen; it may merely represent compensatory cutaneous vasoconstriction.

Changes in Burns.—An extensive burn produces severe *pain* with all its reflex effects; a large volume of protein-rich fluid is lost from the damaged vessels with a resulting decrease in the plasma volume and *anhydræmia*; in a few hours one quarter of the total plasma volume may be lost from the circulation with progressive hæmoconcentration. *Heat loss* is also increased; secondary *infection* may develop. After severe burns the concentration of *histamine* in the blood may rise fourfold in the course of a few days, and this change is paralleled roughly by the gravity of the clinical state, suggesting that it plays a part in producing the symptoms. The importance of the histamine factor must not, however, be overestimated. Products of tissue damage other than histamine may be absorbed.

EXPERIMENTAL AND CLINICAL HYPERTENSION

Experimental Hypertension.—Of the many procedures which have been used experimentally to produce hypertension, two only will be considered here: (i) Section of the sino-aortic nerves; (ii) reduction of the blood flow through the kidneys (renal ischæmia). The latter method is of great clinical interest.

1. Section of the Sino-aortic Nerves.—The chronic persistent hypertension produced by this procedure is described on p. 744.

Clinical hypertension, however, is never the result of denervation by disease of the sino-aortic sensory zones. Strangely enough patients with

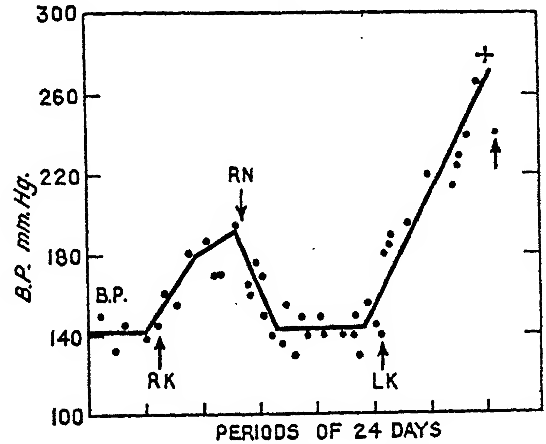


Fig. 206.—Experimental Hypertension from Renal Ischæmia. (Goldblatt, *Ann. int. Med.*, 1937, *11*.)

Record of blood pressure in dog. At RK, constrict right renal artery: progressive rise of blood pressure. At RN, remove right kidney: blood pressure falls again. At LK constrict left renal artery: blood pressure mounts rapidly to 280 mm. and the animal dies.

extensive disease of these zones never seem to suffer from hypertension. In severe clinical hypertension the sino-aortic areas are generally quite normal. The combination of clinical hypertension with intact vasosensory zones means that the factors responsible for the hypertension are so potent that they can completely overcome all the normal regulatory mechanisms.

2. **Renal Ischæmic Hypertension.**[1]—Experimental renal ischæmia produces persistent arterial hypertension. The blood flow through the kidneys is reduced by placing clips on one or both renal arteries and screwing them down to constrict the lumen of the vessel.

(i) If ischæmia is thus induced on *one* side in the *dog* the blood pressure rises steadily for several weeks or months (Fig. 206), but then usually begins to decline, and may ultimately return even to the pre-experimental level.[2]

[1] Braun-Menendez *et al.*, *Renal Hypertension* (trans., Dexter), 1946, Springfield, Illinois.

[2] In the *rat* (p. 356) unilateral renal ischæmia may produce persistent hypertension.

(ii) *Bilateral* renal ischæmia, however, leads to a persistent hypertension[1]; the same result is obtained when one kidney is made ischæmic and the other is completely removed.

(iii) Hypertension is also induced if the aorta is obstructed above the origin of the renal arteries, but not when the obstruction is below, leaving the renal circulation unimpaired.

(iv) In ischæmic hypertension both systolic and diastolic blood pressures are raised and the heart is considerably hypertrophied. The structural changes in the blood vessels are considered on p. 354.

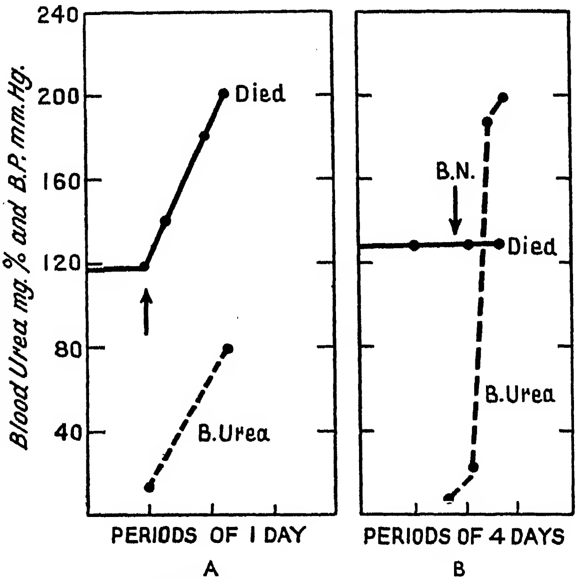


FIG. 207.—Comparison of Effects of Severe Renal Ischæmia and of Nephrectomy on Blood Pressure and Blood Urea. (Goldblatt, *Ann. int. Med.*, 1937, *11.*)

A. At arrow both kidneys were made severely ischæmic; hypertension and urea retention.
B. At arrow (B.N.=bilateral nephrectomy) both kidneys were removed. Urea retention is extreme but the blood pressure remains unchanged.

(v) If the ischæmia is moderate in extent the ischæmic kidneys show little structural change, *e.g.* simple atrophy and perhaps early tubular changes. If the ischæmia is severe there is in addition to the mounting hypertension evidence of grave impairment of renal function ; the usual blood and urinary changes of uræmia develop with a rapidly fatal outcome. The results in an animal experiment are shown in Fig. 207, A.

(vii) A *clinical* case of extensive occlusion of both renal arteries by thrombi s illustrated in Fig. 207. The severe resulting bilateral renal ischæmia caused the blood pressure to rise to 210 mm. Hg (systolic); owing to inadequate renal excretory function the flow or urine was very small and there

[1] When a secondary fall of blood pressure occurs in bilateral experiments it is probably due to the establishment of some degree of collateral circulation which compensates for the diminished blood supply along the main renal arteries.

was marked retention of the non-protein nitrogenous (N.P.N.) constituents of the plasma (see legend to Fig. 208).

Mechanism of Experimental Ischæmic Hypertension.—The hypertension is due to the presence of an ischæmic kidney, for if the affected kidney is removed (in unilateral cases) the blood pressure promptly falls to normal (Fig. 206). It is well known also that hypertension does not follow bilateral

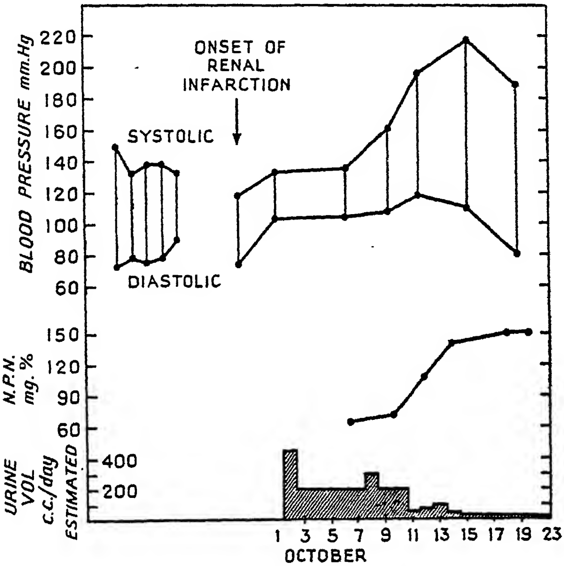


FIG. 208.—Hypertension from Severe Bilateral Renal Ischæmia in Man. (Prinzmetal *et al.*, *J. Amer. med. Assoc.*, 1942, *118*, 44.)

Records from above downwards: systolic and diastolic blood pressure in mm. Hg; non-protein nitrogen of blood in mg-% of N; estimated urinary output in c.c. per day.

A woman aged 50 suffering from auricular fibrillation suddenly developed severe epigastric pain, nausea and dyspnœa; the blood pressure was 115/75. The B.P. gradually rose to 190/110 on the 13th day and the systolic pressure ultimately rose to 210 mm. Hg. The volume of urine diminished till it almost ceased; the non-protein N was 72 mg-% on the 9th day and 148 mg-% on the day before death, which occurred after increasing stupor on the 25th day.

Post-mortem: the entire length of the major right renal artery and the first part of the left renal artery were occupied by thrombotic material; a separate inferior renal artery was patent but its orifice was obstructed by thrombus.

nephrectomy although marked nitrogenous retention occurs; ischæmic hypertension is thus *not* due to any associated rise in the plasma nitrogenous constituents (Fig. 207). The magnitude of the hypertension is related primarily to the degree of renal ischæmia. The ischæmic kidney produces hypertension by *releasing a pressor agent into the circulation.*

EVIDENCE FOR CIRCULATING PRESSOR AGENT.—(i) An *ischæmic* kidney is removed and transplanted into the neck of another completely nephrectomized animal by anastomosing the renal and carotid arteries and the renal and jugular veins respectively. As soon as the circulation is re-established the blood pressure rises rapidly in the recipient by 30–40 mm. Hg in 5–10

minutes (Fig. 209). (In control experiments in which *normal* kidneys are so grafted, hypertensive effects are rarely produced.) If the grafted ischæmic kidney is then transferred into the neck of a second recipient, this animal in its turn shows a rise of blood pressure. The pressure in the first recipient (deprived of its ischæmic kidney) remains high for several hours, after which it gradually returns to normal.

(ii) The venous blood plasma leaving the ischæmic kidney has an intense constrictor action when perfused through the blood vessels of a frog; blood from control kidneys has only a slight activity, or none at all. A perfusate of the ischæmic kidneys from the case recorded in Fig. 208 made immediately after death also contained similar pressor material. The pressor substance is

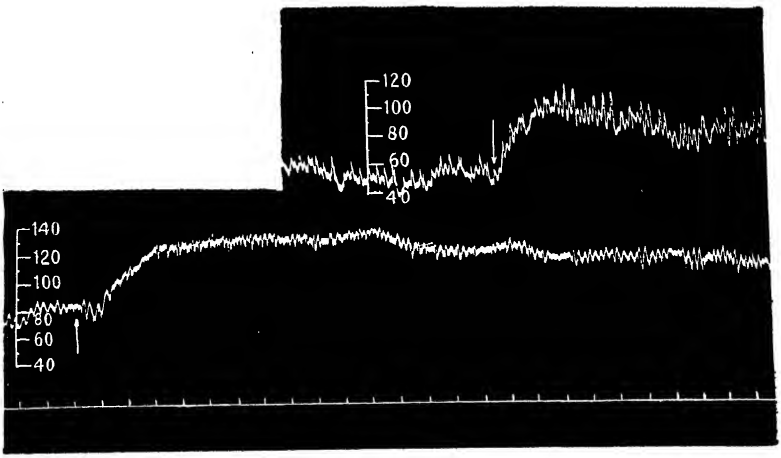


Fig. 209.—Effect of Graft of Ischæmia Kidney on Blood Pressure. (Fasciolo, Houssay and Taquini, *J. Physiol.*, 1938, *94*.)

Record of arterial blood pressure. An ischæmic kidney from a dog with high blood pressure was (at the arrow mark) grafted into the neck, first into one (lower record) and then into another (upper record) chloralosed nephrectomized dog. Note the sustained rise of blood pressure which is produced gradually in each case. Time in minutes.

associated with the *plasma proteins*; it is absent from deproteinized plasma, or from the ultra-filtrate of citrated plasma.

(iii) The vasoconstrictor substance formed by the ischæmic kidney is eliminated or destroyed by healthy kidney tissue. Thus in unilateral renal ischæmia removal of the other, normal, kidney is followed by an aggravation of the pre-existing hypertension. When an ischæmic kidney is grafted into a nephrectomized host (*e.g.* in Fig. 209) it produces a greater rise of blood pressure than in a recipient with intact kidneys. As might be expected, the venous blood leaving the *normal* kidney in a hypertensive animal possesses less pressor activity than the arterial blood reaching it.

(iv) The pressor agent formed by the ischæmic kidney does not act through nervous channels: thus ischæmic hypertension is unaffected by division of the nerve supply to the kidney, section of the splanchnic nerves, division of the ventral roots supplying the abdominal viscera, or even extirpation of

the entire sympathetic chain. The hypertension is not completely abolished even by destruction of the spinal cord.

The pressor substance responsible for the development of ischæmic hypertension is known as *hypertensin* (*angiotonin*). It is formed in the blood as the result of the interaction between an enzyme, *renin*, formed by the ischæmic kidney and a normal plasma globulin (α_2-globulin (p. 136)) called *hypertensinogen*, thus :

$$\underset{\left(\begin{array}{c}\text{normal plasma-}\\\text{protein constituent}\end{array}\right)}{\text{Hypertensinogen}}+\underset{\left(\begin{array}{c}\text{enzyme :}\\\text{protease}\end{array}\right)}{\text{Renin}}\longrightarrow\underset{\begin{array}{c}\text{[Angiotonin]}\\\text{(polypeptide)}\end{array}}{\text{Hypertensin}}$$

The properties of these substances must next be reviewed.

Renin.—(i) Renin is a protein which can be extracted from normal fresh kidney cortex, but not from the medulla. It is present in larger amounts in extracts or perfusates of *ischæmic* kidneys (both experimental and clinical).

(ii) There is *excess renin in the venous blood leaving the ischæmic kidney.*

(iii) On intravenous injection in animals renin raises the arterial blood pressure (Fig. 210) by a peripheral action on the blood vessels. When given by the method of slow continuous intravenous infusion it produces a sustained hypertension.

(iv) Though renin gives rise to vasoconstriction it can be readily shown that it does *not act directly* on the blood vessels. Thus though purified renin markedly raises the blood pressure in the intact animal it has no effect on blood vessels perfused with a Ringer-Locke's solution. But if the perfusion is carried out with *serum*, added renin has its usual constrictor action. If the blood vessels are perfused with recirculated serum, successive injections of renin produce smaller effects ; if fresh serum is now employed, the constrictor effect of renin is restored to its original value. It is clear from all these observations that renin is *not directly constrictor* but *reacts with a serum constituent* which is gradually used up and must either be replaced (in the perfusion circuit) or else given time to be re-formed in the intact animal.

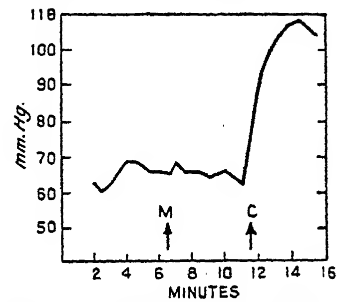


FIG. 210.—Effect of Kidney Extract (Renin) on Arterial Blood Pressure of Unanæsthetized Rabbit. (Pickering and Prinzmetal, *Clin. Sci.* 1938, *3.*)

At M, 1 c.c. of an extract of renal medulla was injected intravenously ; the blood pressure was unchanged. At C, an extract of renal cortex (renin) was injected ; a marked rise of blood pressure occurred.

(v) When renin is incubated with serum the pressor activity of the mixture *increases* initially owing to the appearance of a new substance, a *heat stable polypeptide*, which has the distinctive feature (in which it differs from renin) of constricting blood vessels perfused with Ringer-Locke's solution. This *directly* acting pressor principle is called *hypertensin*. Suitable experiments show that renin interacts with a *globulin* fraction of the serum (called *hypertensinogen*).

(vi) Renin is an enzyme : the yield of hypertensin is increased as the hypertensinogen concentration is increased ; within limits an increase in the renin concentration has the same effect but if the renin level is increased

further the hypertensin yield is diminished. It is concluded that hypertensin is a product of hydrolysis of a globulin (hypertensinogen) by the enzyme renin. Renin is probably a proteolytic enzyme which acts specifically on the blood globulin only; in fact renin may be species specific, the renin extracted from the kidneys of one species not necessarily acting on the blood globulin of another (*e.g.* renin from pig's kidney is inactive on injection in man). [For *physiological* rôle of renin, see p. 353.]

(vii) It is extremely interesting to note that when *pepsin* is incubated with a variety of proteins it also liberates a pressor substance (chemically a polypeptide) resembling hypertensin and called *pepsitensin.*

Hypertensin.—In ischæmic hypertension the presence of *hypertensin can be demonstrated in the circulating blood.*

Injected intravenously in *animals*, hypertensin produces a sharp transient rise of blood pressure resembling that produced by adrenaline and due to a peripheral constrictor action on the blood vessels. It differs from adrenaline in showing none of the drug antagonisms of the latter and in being motor to the isolated intestine (cf. p. 724). Slight tachyphylaxis develops after a long series of injections, *i.e.* the pressor responses progressively decline in magnitude. There is some evidence that the blood may contain an agent which *potentiates* the action of injected hypertensin (hypertensin-" *activator* ") and which is slowly used up.

Action of Hypertensin in Man.[1]—(1) *Action on Circulation.*—(i) When hypertensin is injected or infused intravenously in normal man it raises both systolic and diastolic blood pressure to an extent which depends on the dosage used (Fig. 211). It *constricts the peripheral blood vessels*: (*a*) injected intradermally it produces blanching of the skin; (*b*) injected intra-arterially it produces vasoconstriction in the muscle mass supplied by the artery; (*c*) injected intravenously the blood flow through the limbs decreases.

(ii) The effects on the *heart* are complex; (*a*) there is marked cardiac slowing which is abolished by atropine (indicating vagal involvement); (*b*) there is usually evidence of heart failure (Fig. 211), *e.g.* diminished cardiac output, rise of venous pressure, engorgement of the lung vessels (as shown by the decreased vital capacity), increased circulation time, and an increase in the size of the heart. The arterial hypertension produced by injection of large doses of hypertensin in man is thus due to vasoconstriction and still occurs even if there is coincident heart failure.

(iii) The cardiac changes just described are those generally observed when sudden marked vasoconstriction and hypertension are produced experimentally and the load proves temporarily excessive for the heart to cope with. It is likely that if smaller amounts of hypertensin were infused steadily for much longer periods compensatory cardiac hypertrophy would take place, the cardiac efficiency would increase, and the heart would sustain the hypertension without signs of failure such as venous engorgement.

(iv) Although intravenously injected hypertensin produces some constriction of the skin vessels as shown by a fall of skin temperature, the contraction of these vessels can be readily overcome by chemical agencies, *e.g.* by reactive hyperæmia.

(2) *Action on Kidney.*—Hypertensin produces important changes in renal function (Fig. 212).

[1] Bradley and Parker, *J. clin. Investig.*, 1941, *20*, 715; Wilkins and Duncan, *ibid.*, 721.

(i) The renal *blood flow* is markedly reduced to a greater extent than would be expected from the impairment of the heart's action ; it is therefore the result of constriction of the renal blood vessels.

(iii) *Glomerular filtration rate* is also cut down, but to a less extent proportionately than the renal blood flow.

(iii) The *filtration fraction* (glomerular filtrate/renal plasma flow) rises,

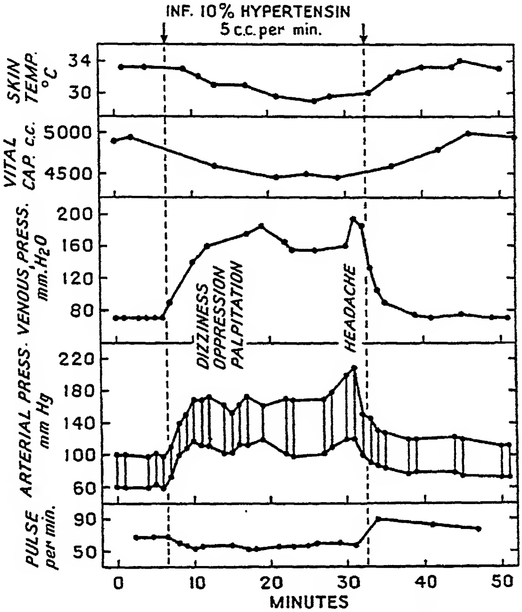


FIG. 211.—Effects of Intravenous Infusion of Hypertensin (Angiotonin) on the Circulation in a Normal Man. (Wilkins and Duncan, *J. clin. Investig.*, 1941, *20*, 729.)

In arterial pressure records, upper line = systolic, lower line = diastolic, vertical lines = pulse pressure. At 6½ minutes (first interrupted vertical line) commence infusion of a 10% solution of hypertensin, at rate of 5 c.c. per minute ; infusion stopped at 32½ minutes (second interrupted vertical line).

Note : (i) Fall of skin temperature = constriction of skin arterioles.
(ii) Decrease in vital capacity = pulmonary congestion (cf. p. 307).
(iii) Rise of venous pressure = heart failure.
(iv) Rise of systolic, diastolic, and pulse pressures (due to vasoconstriction).
(v) Slowing of heart.

The symptoms are noted on the chart.

e.g. from 0·2 to 0·3, *i.e.* a greater proportion of the plasma flow is filtered out into Bowman's capsule. This finding is attributed to raised intra-glomerular pressure ; it must therefore be supposed that the efferent glomerular arterioles are more markedly constricted than the afferent arterioles (cf. pp. 25, 27).

(iv) The *concentrating power* of the kidneys is decreased, *i.e.* less water absorption takes place in the tubules.

(3) The *subjective* symptoms consist of palpitation, a feeling of substernal oppression, dizziness, headache, and nausea.

DESTRUCTION OF HYPERTENSIN.—Hypertensin is readily destroyed by incubation with tissue extracts or *serum* which contain a proteolytic enzyme,

hypertensinase ; this enzyme probably breaks down the hypertensin polypeptide into inactive amino-acids. This enzyme is possibly formed by the normal kidney.

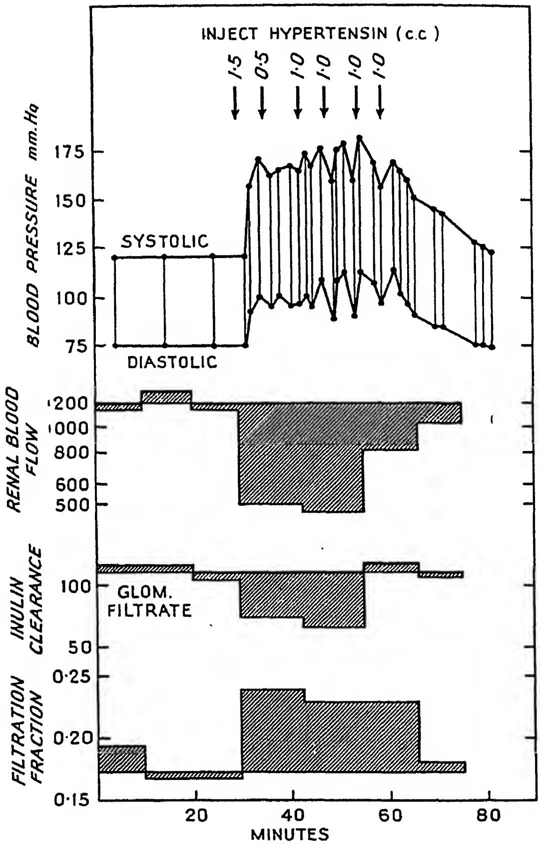


FIG. 212.—Effect of Infusion of Hypertensin (Angiotonin) on Renal Function in Man. (Page, *J. Mt. Sinai Hosp.*, 1941, *8*, 14.)

Records from above downwards : systolic and diastolic blood pressure ; renal blood flow (diodrast clearance) in c.c. per minute ; glomerular filtration rate (inulin clearance) in c.c. per minute ; filtration fraction, *i.e.*, glomerular filtrate/renal plasma flow ; arrows indicating volume of injected hypertensin solution. Time in minutes.

Note : (i) Rise of arterial blood pressure.
(ii) Decrease in renal blood flow from 1200 c.c. to under 500 c.c. per minute.
(iii) Reduction in glomerular filtrate from 150 to 75 c.c.
(iv) Rise in filtration fraction to about 0·3.

The lower three records are plotted semi-logarithmically. The horizontal line on each of these records indicates mean normal values.

(i) The *nephrectomized* animal responds to an injection of hypertensin with a greater and more sustained rise of blood pressure than the normal animal ; this result may be due to the presence of less hypertensinase.

(ii) Similarly the grafted ischæmic kidney produces a greater rise of blood pressure in the nephrectomized animal (p. 348).

(iii) In unilateral ischæmia, removal of the normal kidney aggravates the hypertension.

SUMMARY OF MECHANISM OF ISCHÆMIC HYPERTENSION.—Ischæmia of the kidney disturbs its metabolism leading to the excessive discharge into the circulation of the proteolytic enzyme renin which acts on a plasma α_2-globulin, hypertensinogen, to form the polypeptide hypertensin. The action of hypertensin may be potentiated by an "activator" in the plasma. Hypertensin is normally destroyed by hypertensinase which is provided mainly by the kidney.

PHYSIOLOGICAL RÔLE OF RENIN.—It has been suggested that when the arterial blood pressure falls, renin may be *normally* secreted to produce compensatory vasoconstriction. In support of this hypothesis two observations may be quoted: (i) during severe hæmorrhage the renin content of the venous blood from the kidney rises; (ii) after nephrectomy animals succumb to a smaller loss of blood than do intact animals. Nothing is known about the way in which such raised renin secretion is brought about; it may be the direct result of the reduced renal blood flow.

Little is known about the effects of injection of renin in *man*. Renin of animal origin is inactive, and potent preparations of human renin are not yet available.

Clinical Hypertension.—The clinician uses the term hypertension to refer to a persistent elevation of both systolic and diastolic blood pressure; the upper normal limit of pressure in the adult is 160 mm. Hg systolic and 100 mm. Hg diastolic; in doubtful cases more significance should be attached to the raised *diastolic* blood pressure. Most cases of hypertension fall into one of the following groups:

(i) *Essential* ("*primary*") *hypertension:* (*a*) benign; (*b*) malignant.

(ii) *Secondary hypertension:* (*a*) renal; (*b*) endocrine.

Essential Hypertension.—The condition is so named because its causation is not yet established.

(1) BENIGN FORM.—In the early stages the hypertension is moderate. The blood pressure, especially the systolic, fluctuates considerably: during sleep or emotional and physical rest, the pressure may be normal; in states of stress the pressure rises to pathological levels. Later the hypertension becomes "fixed" in the pathological range and cannot be reduced to normal by rest or sedatives like the barbiturates. There is compensatory cardiac hypertrophy; the walls of the small arteries and arterioles become thickened (p. 354); renal changes appear, *e.g.* an increased volume of night urine, albuminuria or slight hæmaturia. After a period which may vary from a few years to 20 years death occurs from heart failure, vascular accidents (hæmorrhage or thrombosis), or renal failure.

(2) MALIGNANT FORM.—The condition is so named because death occurs within 6 months to 2 years of its first recognition. The blood pressure is much higher than in the benign form; the peripheral vascular changes are severe and degenerative (p. 354) and readily observed in the retinal vessels; papillœdema is frequently present. Severe renal failure is common.

It is possible, but by no means certain, that the benign and malignant conditions are not diseases of independent ætiology, but are different grades of severity of the same disease.

The *functional* changes found in essential hypertension will now be described.

common in the retinal and cerebral vessels but not in the vessels to skeletal muscle, skin, and lungs. The degenerative changes are of the following types :

(i) *Diffuse Fibrosis.*—This is the common change in *benign* hypertension. There is *fibroblastic* replacement of the muscular fibres of the media and

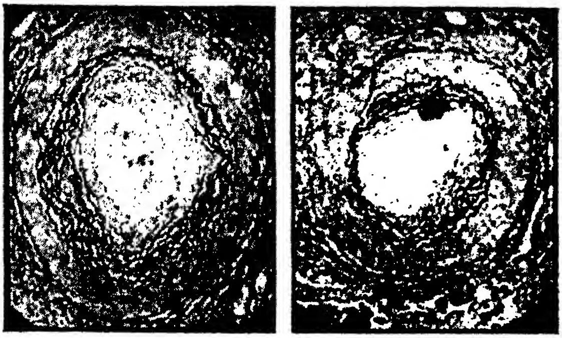
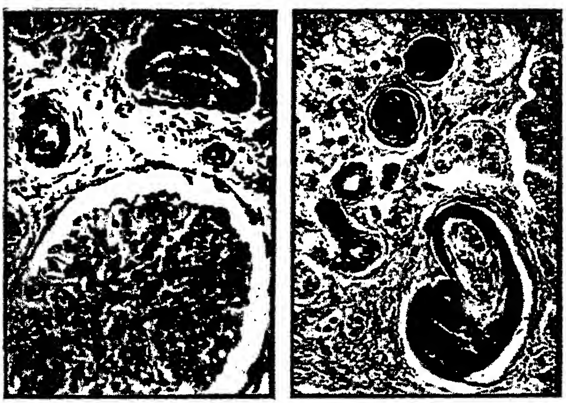


FIG. 213.—Hypertrophic Changes in Interlobular Artery of Kidney in Benign Essential Hypertension.

The interlobular artery has been cut at two points along its course. The section is stained for elastic tissue. Note the great uniform increase of the subintimal connective tissue and the numerous newly formed layers of elastic tissue. (Hadfield and Garrod, *Recent Advances in Pathology*, 5th edn., J. & A. Churchill, London, 1947.)

reduction of the lumen of the vessel by proliferation of the subintimal connective tissue.

(ii) *Fibrinoid Degeneration.*—This process affects many visceral vessels but it is noteworthy that in nearly every case the *afferent* glomerular arteriole is affected (Fig. 214). There is enormous swelling of the vessel wall owing to



A B

FIG. 214.—Arteriolar Degeneration in Essential Hypertension.

A. Kidney. Note fibrinoid degeneration of afferent glomerular arterioles.
B. Adrenal cortex. Note great thickening and fibrinoid degeneration of arterioles of varying size.
(Hadfield and Garrod, *Recent Advances in Pathology*, 5th edn., J. & A. Churchill, London, 1947.)

the appearance at first under the intima of hyaline eosinophil structureless material which progressively replaces the muscle coat and gradually obliterates the lumen. The degeneration begins at the proximal end of the afferent glomerular arteriole and gradually spreads distally, finally involving the glomerular tuft itself (Fig. 215).[1]

(iii) *Acute Arteriolar Necrosis.*—This is the typical lesion of *malignant* hypertension. It occurs generally in vessels that have *not* undergone initial hypertrophy and which may still be in the fibrinoid stage. Hæmorrhage occurs into the vessel wall which simultaneously undergoes necrosis.

RELATION OF HYPERTENSION TO STRUCTURAL CHANGES IN VESSELS.—There is persuasive evidence that both in experimental ischæmic hypertension

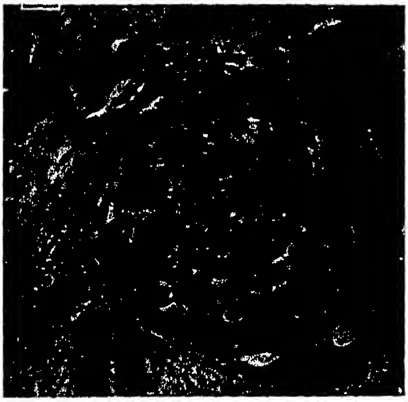


FIG. 215.—Vascular Changes in Ischæmic Rena Hypertension (in the rat). (Wilson and Byrom, *Lancet*, 1939, i, 136.)

Rat kidney. Necrosed glomerulus with swelling and proliferation of capsular epithelium; fibrinod necrosis of afferent arteriole. ×260.

and in clinical essential hypertension the *rise of blood pressure is primary* and the *structural changes are secondary.*

(i) Thus the hypertension may set in within a few hours of the production of renal ischæmia.

(ii) In rats, *unilateral* renal ischæmia often produces a persistent hypertension. In the ischæmic kidney which is protected (by the clamp on the renal artery) from the full strain of the hypertension, the structural changes are slight and consist of simple tubular atrophy. In the kidney with the free circulation which is consequently fully exposed to the hypertension the typical vascular changes of severe clinical hypertension appear, involving the interlobular arteries and the afferent glomerular arterioles.

(iii) In experimental rats with rapidly developing severe hypertension, and especially when the blood pressure fluctuations are large, the vascular changes resemble those seen in malignant hypertension, *i.e.* they are degenerative and mainly fibrinoid or necrotic.

[1] It is noteworthy that though pathologically the *afferent* glomerular vessel is principally involved, the clinical evidence suggests that the *efferent* glomerular vessel is more constricted.

(iv) In rats in which the hypertension developed gradually and was more moderate, medial hypertrophy of the arterioles was the predominant lesion. The structural changes are independent of the presence or absence of renal failure.

(v) The pathological evidence suggests that in clinical hypertension likewise the rise of blood pressure is primary and the structural vascular changes are secondary.

CAUSE OF ESSENTIAL HYPERTENSION.—If the above reasoning is correct, the increased peripheral resistance which produces the hypertension must *initially* be due to an *active* contraction of the walls of the arterioles. Such vasoconstriction may be due to sympathetic overaction or to the presence of a circulating pressor agent.

(1) *Sympathetic Overaction.*—(i) In favour of this view is the marked influence of emotional tension in producing or aggravating the hypertension.

(ii) The blood pressure is temporarily reduced by full doses of the barbiturates; these drugs are known to depress the hypothalamic nuclei.

(iii) Section of the splanchnic nerve and the lateral sympathetic chain in selected cases markedly lowers the blood pressure for months or years. It is doubtful, however, whether lasting effects are commonly produced by these operations.

(2) *Circulating Agent.*—In view of the many resemblances between essential and experimental ischæmic hypertension, it has been suggested that in essential hypertension a circulating pressor agent is involved. But in the early stages of the clinical condition there are no anatomical changes in the renal vessels which could produce ischæmia; if there is *initial* renal ischæmia it must be due to *spasm* of the vessel walls due either to sympathetic overaction or to some *unidentified* pressor substance. Once, however, the hypertension is established (by whatever means) the raised blood pressure *secondarily* produces structural changes in the blood vessels including especially those of the kidney. As the small vessels become thickened and narrowed a permanent increase in peripheral resistance occurs which is unamenable to treatment. The renal changes produce renal ischæmia with the possible release of excess renin and the consequent formation of hypertensin. A *hypertensive vicious circle* might thus be established: hypertensin produces renal ischæmia; the renal ischæmia results in increased circulating hypertensin and greater hypertension; the increased hypertension produces more severe renal ischæmia and thus the hypertension steadily mounts and the renal damage becomes progressively more grave. This plausible and attractive hypothesis suffers from the defect that neither excess renin nor hypertensin have yet been demonstrated in the arterial or venous blood in essential hypertension whether renal failure was absent or present.

Secondary Hypertension.—(1) ENDOCRINE GROUP.—Hypertension occurs in tumours of the adrenal medulla (*phæochromocytoma*, p. 734) and in some cases of *Cushing's syndrome* (*hypercortism*).

(2) RENAL GROUP.—In this group of cases which constitute about 10% of all cases of hypertension there is clear evidence that the renal lesion is *primary*, *i.e.* it undoubtedly precedes the development of hypertension and is supposed to be the causal factor.

(i) *Glomerulo-Nephritis.*—In Type 1 nephritis (p. 75) the inflammatory process markedly affects the renal vessels and produces gross renal

ischæmia. A hypertensive vicious circle may thus be set up leading to mounting hypertension and progressive renal destruction as explained on p. 357.

(ii) "*Surgical*" *Kidney Disease.*—Cases have been recorded in which unilateral or bilateral hydro- or pyelonephrosis was associated with hypertension: removal of the affected organ, when the disease was *one-sided*, led to complete return of the circulation to normal. It is likely that the disease, in such cases, mainly affected the renal blood supply and led to excessive secretion of renin and consequently to excessive formation of hypertensin.[1] These cases are uncommon.

DISORDERS OF PERIPHERAL CIRCULATION

Raynaud's Disease.[2]—This condition affects the parts of the body that are particularly influenced by exposure to cold, most commonly the hands, and less often the feet, ears, or nose. If patients with this disease are exposed to external cold when the skin temperature is also low, attacks develop in which the skin (usually of the hands) becomes discoloured and numb. The sequence of events is as follows: The digital arteries become completely closed, first at the finger-tips and later at the bases; the circulation in the fingers consequently ceases. This vascular occlusion leads after a short delay to two secondary events: (i) numbness, from loss of function of the cutaneous sensory nerve endings; (ii) closure of the capillaries and venules, and therefore blanching of the skin. When recovery sets in, the skin becomes first violet and later red in colour. The latter flush is due to vascular dilatation, *i.e.* the reactive hyperæmia which always follows on vascular occlusion, however induced (p. 329). Small areas of necrosis may appear at the finger-tips after repeated attacks.

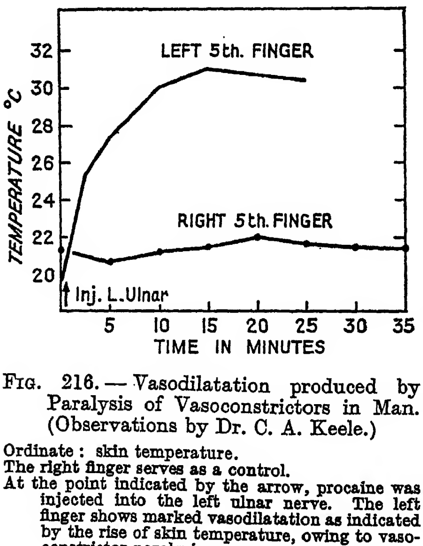


FIG. 216.—Vasodilatation produced by Paralysis of Vasoconstrictors in Man. (Observations by Dr. C. A. Keele.)

Ordinate: skin temperature.
The right finger serves as a control.
At the point indicated by the arrow, procaine was injected into the left ulnar nerve. The left finger shows marked vasodilatation as indicated by the rise of skin temperature, owing to vasoconstrictor paralysis.

The mechanism involved in the attacks must now be considered. They are not due simply to excessive vasoconstrictor discharges along the sympathetic nerves; thus paralysis of the vasoconstrictors with procaine during an attack does not relieve the spasm, or only slightly, and after considerable delay; in normal subjects, on the other hand, even when the skin is cold, this procedure leads to immediate vasodilatation (Fig. 216). Attacks of Raynaud's

[1] Weiss and Parker, *Medicine*, 1939, *18*, 221.

[2] Lewis, *Vascular Disorders of Limbs*, London, 1936. Lewis and Pickering, *Clin. Sci.*, 1934, *1*, 327; 1938, *3*, 287, 321.

disease can occur in severe cases even after sympathectomy. Lewis attributes the disease primarily to an *abnormality of the small peripheral vessels* of the order of size of the digital arteries. An attack can be brought on by local cooling of the digital arteries, the spasm being confined to the part which has been cooled. (Normal vessels never react in this way.)

When the body is exposed to cold the surface vessels are constricted partly by a direct action and partly reflexly through the sympathetic vasomotor nerves (p. 328). In normal subjects the superficial vessels are markedly narrowed; in Raynaud's disease, however, owing to the additional factor of the *local abnormality,* the vessels become completely closed. But if the hand of a patient is kept warm, general exposure to cold which can now act only reflexly on the blood vessels does not induce an attack. Following sympathectomy the blood vessels in the limb become dilated and perhaps more resistant to the effects of local cold, and of course they cannot be influenced reflexly by vasoconstrictor nerves; but there is not complete immunity to attacks (cf. p. 361).

Sympathectomy for Vascular Disorders in Man.[1]—The immediate vascular changes following sympathectomy are the same whether the operation is performed on the preganglionic fibres, the sympathetic ganglia or the postganglionic fibres, so long as all the fibres supplying the part are severed. The end results depend: (i) on whether regeneration of fibres takes place; (ii) on the natural course of the disease for which the operation is carried out.

Upper Limb.—The operation of choice is a preganglionectomy. The lateral sympathetic chain is cut between the third and fourth thoracic ganglia; this divides all the preganglionic fibres destined for the arm which leave the spinal cord below Th3 (cf. p. 709). To cut any possible preganglionic outflow in Th2 and Th3 the dorsal and ventral roots are severed and a piece of the 2nd and 3rd intercostal nerves cut out. Finally the distal end of the sympathetic chain is turned up and stitched to muscle to prevent regeneration.[2]

Lower Limb.—The operation is mainly a preganglionectomy: the second, third, and fourth lumbar ganglia and the intervening sympathetic

[1] Adson and Brown, *Arch. Neurol. Psychiat.*, 1929, *22*, 322. Lewis and Landis, *Heart*, 1930, *15*, 151.

[2] The older operation was mainly a ganglionectomy: it consisted of removal of the stellate ganglion, *i.e.* the first thoracic and inferior cervical ganglia, and division of the communicating branch to the first thoracic nerve from the second which often receives a grey ramus from the second thoracic sympathetic ganglion. Such an operation destroys most of the excitor cells supplying the upper limb and divides the connector fibres passing to the middle cervical ganglion. As the excitor cells are extensively removed the risk of regeneration is minimal. The operation has two disadvantages, one real and one speculative.

(i) The real is due to dividing the sympathetic fibres to the head and neck and so producing Horner's syndrome (p. 708), especially drooping of the upper lid and constriction of the pupil.

(ii) The speculative is based on the fact that the degeneration of the postganglionic fibres and terminals which follows ganglionectomy increases the sensitivity of the denervated parts to the action of injected or naturally secreted adrenaline (p. 729). Therefore any adrenaline which is poured out in states of emotional stress, physical exertion, or exposure to cold might produce an undue degree of vasoconstriction which would tend to annul the dilator effects of the sympathectomy.

trunk are removed. [The preganglionic outflow to the lower limb is from the lower thoracic and upper lumbar roots (p. 709).]

RESULTS OF SYMPATHECTOMY.—After a successful operation vasodilatation immediately sets in, as shown by a rise of skin temperature, flushing of the skin, and increased loss of heat; the superficial veins are also relaxed. This occurs in subjects with normal blood vessels (*e.g.* patients with spastic paralysis in whom the sympathetic was divided under the erroneous impression that it would decrease skeletal muscle tone), and in appropriately selected cases of vascular disorder. Reflex sweating and pilomotor reactions are abolished. The dilatation may persist for months or years. After a time

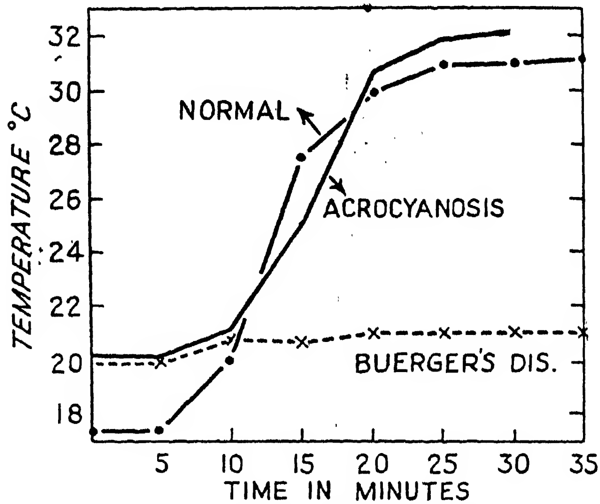


FIG. 217.—Vascular Response of Skin of Fingers to Heating of Body. (Observations by Dr. C. A. Keele.)

Ordinate: skin temperature.
The body is placed in a heating chamber, but the arms are excluded. In the normal person and in the case of acrocyanosis the rise of body temperature produces marked vasodilatation in the fingers as shown by the progressive rise of skin temperature (cf. Fig. 198). No vasodilatation (no rise of skin temperature) occurred in this case of Buerger's disease (thromboangiitis obliterans) (cf. p. 361).

the skin becomes *paler* on the operated side because of contraction of the small superficial vessels; Lewis suggests this is due to the increased blood flow through the skin preventing the natural local vasodilator substances from accumulating to their normal concentration.

INDICATIONS FOR SYMPATHECTOMY.—Before sympathectomy is performed for the relief of disorders of the peripheral circulation, certain preliminary tests must be carried out to determine what the immediate outcome of the operation is likely to be.[1]

(i) Inject procaine into a suitable nerve (*e.g.* the ulnar or peroneal) to paralyse the vasoconstrictor fibres. In the normal person marked vasodilatation and a great rise of skin temperature (*e.g.* from 20° to 30° C.) occur in the corresponding fingers or toes (Fig. 216).

[1] Lewis and Pickering, *Heart*, 1931, *16*, 33.

(ii) Spinal anæsthesia may be induced up to the level of the mid-thoracic cord to paralyse the vasomotor fibres to the legs. Similar results to those described in (i) *supra* are normally obtained in the toes.

(iii) The patient is enclosed in a suitable chamber, the temperature in which can be raised; the limbs to be investigated are kept *outside* the chamber. As the chamber temperature rises, the body temperature goes up. The heated blood acts on the vasomotor centre diminishing the vasoconstrictor discharge to the excluded limbs; afferent impulses from the heated skin also act reflexly in the same way (p. 327). Skin temperature in the excluded limbs (especially in the fingers and toes) rises owing to the resulting decreased vasomotor tone; there may perhaps also be active stimulation of the sympathetic vasodilators (Fig. 217).

(iv) In cases where the decreased blood supply to *muscles* is under analysis the blood flow to the calf or forearm (using the technique of venous occlusion plethysmography (p. 305) should be examined before and after blocking the vasoconstrictor nerves, to see whether any increase takes place.

Unless the above positive reactions are obtained, sympathectomy will not produce even immediate or transient benefit and is entirely contra-indicated. Positive reactions (*i.e.* vasodilatation) are elicited in *acrocyanosis* (Fig. 217), and in patients with *Raynaud's disease between the attacks.*

Results of Sympathectomy.—(1) *Raynaud's Disease.*—In this disease sympathectomy often produces temporary benefit, but the lasting improvement is less spectacular. Various explanations have been offered for these somewhat disappointing results. Lewis suggests that the disease is due to a primary disorder of the vessel wall of a progressive character; under these circumstances the most that can be hoped for is that the frequency and intensity of the attacks of vascular spasm will be diminished. In some cases regeneration of the sympathetic fibres (as shown by return of reflex vasodilatation and sweating) is a factor.

(2) *Thromboangiitis Obliterans.*[1]—In this disease, organic changes in the main arteries of the legs lead to partial occlusion of the lumen and a decreased blood flow to the distal part of the limbs. The patient complains mainly of the results of relative ischæmia of the muscles during walking (*intermittent claudication*, p. 751); in addition, there may be trophic changes in the skin from inadequate blood supply. Sympathectomy cannot increase the lumen of the affected large vessels, but it may help in another way. The blood supply to the distal parts of the limb is kept up in this disease mainly by way of a collateral circulation from the normal vessels above the block. Sympathectomy by dilating arterioles might improve such collateral circulation; it also relaxes the skin vessels and to a less extent the muscle vessels. Examination by methods (i) to (iv) *supra* will show how much improvement in skin (Fig. 217) and especially in *muscle* flow is likely to follow sympathectomy. A few selected cases have benefited from the operation.[2]

[1] Also called *Buerger's disease.*
[2] Cf. Ascroft, *Brit. J. Surg.*, 1937, *24*, 787; *Brit. med. J.*, 1937, ii, 173.

IV

RESPIRATION[1]

GENERAL CONSIDERATIONS[2]

Subdivisions of Lung Volume[3] (Fig. 268).—(1) TIDAL AIR (*resting tidal volume*) is the volume of air breathed in or out of the lungs with every quiet respiration. The normal range is 350–500 c.c.

(2) COMPLEMENTARY AIR (*inspiratory capacity*) is the volume of air that can be taken in after a normal expiration by a maximum inspiratory effort; it includes the tidal air. Its volume is 2000 to over 3500 c.c.

(3) RESERVE AIR (*expiratory reserve volume*) is the largest volume of air that can be expired after a normal expiration. The average is about 1500 c.c.

(4) **Vital Capacity**[4] is determined by measuring the maximum volume of air that can be expired after a maximal inspiration; the range is 3500 to over 5000 c.c. It is the deepest breath that we can take, and is about eight times as great as the normal-sized breath. The vital capacity varies with the size of the individual and the use to which he has put his respiratory apparatus. It is related to the *surface area*; it amounts to 2·6 litres per square metre of surface in a male, and 2·1 litres per square metre in a female (on the average). Most normal people have a vital capacity within 10% of these values. Vital capacity is affected by the following factors:

(i) The *development and power of the respiratory muscles*: it is increased in athletes, especially those who use their respiratory apparatus vigorously, *e.g.* swimmers or divers; it is increased by practice; it is decreased in general physical weakness (*e.g.* by 20–30%) and in older people.

(ii) The *integrity of the respiratory centre*, its descending *tracts*, the related *spinal motor neurones* and the *respiratory muscles*: the vital capacity is therefore decreased when the respiratory centre is depressed by narcotics; in injuries to the spinal cord, *e.g.* trans-section in the cervical or thoracic region; destruction of spinal motor neurones (cells or efferent fibres), *e.g.* in poliomyelitis, ventral nerve root lesions, avulsion of the phrenic nerves; and in disorders of the muscles, *e.g.* myasthenia gravis or therapeutic curarization.

(iii) *Patency* of respiratory pathways: the vital capacity is decreased by obstruction of the larynx, trachea, bronchi or bronchioles.

(iv) The *volume of blood* in the lungs (cf. p. 307): vital capacity is greater in the standing than in the sitting or recumbent position partly because of the decreased pulmonary blood volume. In congestive heart failure the vital

[1] See Haldane and Priestley, *Respiration*, New Haven, new edn., 1935. For methods, see Douglas and Priestley, *Human Physiology*, Oxford, 3rd edn., 1948; Peters and Van Slyke, *Quantitative Clinical Chemistry, Methods*, London, 1932.

[2] Miller, *The Lung*, Springfield, Ill., 1937.

[3] For terminology, see *Fed. Proc.*, 1950, *9*, 602.

[4] Hurtado and Boller, *J. clin. Investig.*, 1933, *12*, 793.

capacity is decreased proportionately to the degree of pulmonary congestion (and œdema). (Cf. also Fig. 211.)

(v) *Space-occupying lesions* of the thorax, *e.g.* pleural effusion, closed pneumothorax (cf. p. 371) or tumours, which reduce to some extent the maximal air entry into the lungs. In patients with *open* pneumothorax the vital capacity is further reduced because air enters via the opening in the chest wall as well as by the normal path (p. 368).

(vi) *Disease of the lung substance, e.g.* fibrosis or pulmonary tuberculosis, which interfere with its expansion. There is also a marked reduction of vital capacity in emphysema. The explanation is not clear : it has been attributed to decreased lung elasticity, which impairs the efficiency of expiration, or to fixation of the costo-sternal junctions, which interferes with the respiratory movements of the ribs (p. 459).

(vii) *Œdema* of the lungs, which results from many causes (p. 116) decreases vital capacity partly because the fluid replaces air and partly because the respired air and fluid are beaten up into a froth which mechanically hinders the flow of air to and from the lungs.

(viii) Abnormalities of the *bony thoracic cage* : for example, deformities of the vertebral column, sternum or ribs, or fixation of the costo-sternal or costo-vertebral joints reduce vital capacity.

(ix) The respiratory movements and vital capacity may be reduced by pain (*e.g.* of pleurisy or of some abdominal disorder), or hampered mechanically by an increase in the intra-abdominal contents due, *e.g.* to pregnancy, tumour, ascites or flatus.

(x) In *exophthalmic goitre* the vital capacity is reduced for an unknown reason.

Relation of vital capacity to dyspnœa.—As the depth of breathing approaches the vital capacity subjective feelings of discomfort (*dyspnœa*) develop. Patients with reduced vital capacity consequently become breathless on exertion more readily than normal people (cf. p. 455).

On an *average* there is a fairly satisfactory relationship between the clinical estimate of respiratory efficiency and the degree of reduction of the vital capacity. But some patients who have extensive pulmonary disease may show a vital capacity which is within normal limits ; significance can therefore only be attached to a *low* value for the vital capacity.

(5) Residual Air (or Volume) is the air which remains in the lungs after a maximal expiration ; it can only be expelled by opening the chest widely and allowing the lungs to collapse. The average volume in adults is 1500 c.c. The *functional residual air* (or *capacity*) is the volume of air left in the lungs after a normal expiration. It is roughly equal in volume to that of the alveolar air (*infra*).

(6) Total Lung Volume (or *capacity*) is the sum of the residual air and the vital capacity ; on an average it is 5500 c.c.

(7) **Alveolar Air.**—As the bronchioles branch and become smaller their *columnar ciliated* epithelium becomes modified into a *cubical* epithelium The terminal branches are called *respiratory* bronchioles (Fig. 218), because in places the normal bronchial wall is replaced by patches of thin flattened cells through which gaseous interchanges can take place. Each respiratory bronchiole leads into an expansion, the *vestibule*, from which arise several passages, the *atria*. Each atrium leads into two or three *air-sacs* (*infundibula*) ; the walls of the air-sacs are studded with minute evaginations, the

alveoli. The alveoli are lined by large, thin, flattened cells, but in places clumps of cubical granular cells are seen. External to the lining epithelium are yellow elastic fibres and numerous thin-walled blood vessels. The term, alveolar air, does *not* refer to the air which is present in the anatomical alveoli, but is used to describe the air in the *depths of the lung which is more or less in contact with the respiratory epithelium, and can thus carry out gaseous interchanges with the blood.* Alveolar air is a physiological and not an anatomical entity. It is approximately equivalent to the "functional residual air," *i.e.* the sum of the reserve air and the residual air; the average volume is about 3000 c.c.

A sample of alveolar air can be collected by the method of Haldane and Priestley. A rapid *maximal expiration* is made down a narrow tube about 4 feet long, and a sample is obtained of the air which is expelled last from the lungs. By this means the air in the mouth, pharynx, and air passages (the dead-space air, *infra*) is driven out and alveolar air from the depths of the lung is collected; it is analysed with the Haldane gas analysis apparatus. The composition of alveolar air is: CO_2 4·7–6·4% (average 5·6%), oxygen about 14%, nitrogen 80%; it is saturated with water vapour at body temperature. The average *tension* (or *pressure*) of alveolar CO_2 is 40 mm. Hg; that of alveolar oxygen is 100 mm. Hg.

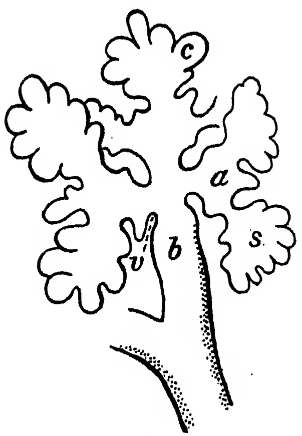


FIG. 218.—Diagram of Termination of Respiratory Bronchiole.
b=Respiratory bronchiole; *v*=Vestibule; *a*=Atrium; *s*=Air-sac *c*=Alveolus. (After Miller.)

The tension of the gases in alveolar air is calculated from the percentage composition thus: the total pressure exerted by all the gases in alveolar air is equal to atmospheric pressure (*e.g.* 760 mm. Hg). Alveolar air is saturated with water vapour at 37° C. which exerts a pressure of 47 mm. Hg; the pressure of the dry gases (CO_2, O_2, N_2) is consequently $760-47=713$ mm. The average tension of CO_2 is therefore $\frac{5·6}{100}\times 713=40$ mm.; that of O_2 is $\frac{14}{100}\times 713=100$ mm.

The breathing is normally adjusted to maintain a fairly constant composition of the alveolar air, because the *alveolar air regulates the tension of the gases in the arterial blood.* Enough fresh air has to be introduced into the alveoli to replace the oxygen taken up by the blood, and to get rid of the CO_2 which has come out of the blood. The venous blood which arrives at the lungs gives off CO_2 and takes up oxygen until it comes into equilibrium with the air in the alveoli. The tensions of the gases in arterial blood are therefore normally about equal to those in alveolar air. *The arterial gas tensions can thus be approximately deduced if the alveolar gas tensions are known,* so long as the pulmonary epithelium is normal in character.

(8) DEAD-SPACE AIR is found in the air passages, *i.e.* the nasopharynx, trachea, bronchi. It does not come into contact with the pulmonary

epithelium and carries out no interchanges with the blood. Expired air is a mixture of dead-space air (which is identical in composition with inspired air) and of alveolar air. The volume of the dead space can be calculated if the composition of inspired, expired, and alveolar air and the volume of the tidal air are known.[1]

The formula usually used is:

$$\text{Dead Space} = \frac{\text{Tidal air} \times (CO_2\%\text{ in alveolar air} - CO_2\%\text{ in expired air})}{CO_2\%\text{ in alveolar air.}}$$

It is probable that the *anatomical* dead space—in other words, the actual capacity of the respiratory passages—alters little in different clinical conditions and is affected only by constriction or dilatation of the bronchioles. Calculated in the manner described above, the dead space usually amounts to about 150 c.c. and constitutes 25–30% of the tidal air.

Air Movements during Respiration.—During inspiration the size of the thoracic cavity is increased. The air in the depths of the lungs follows the outward movement of the chest wall. As the capacity of the lungs is increased, first the dead-space air and, later, the fresh air from outside enter the depths of the lungs. One must imagine an interface between the air previously present in the alveoli, and that which has newly entered, across which gaseous interchange occurs. The newly inspired air does not come directly into contact with the pulmonary epithelium. The inspired air during quiet breathing amounts to only one-seventh of the air permanently present in the lungs. *The alteration produced in the alveolar air by the fresh air breathed just compensates for the interchanges which are taking place between the alveolar air and the pulmonary blood.* The efficiency with which the tidal air "washes" the alveolar air depends to a considerable degree on the amplitude of breathing. If the depth of the breathing were reduced to the size of the dead space no ventilation of alveoli would take place at all.

Laws of Gases.[2]—(1) At constant temperature, the volume (V) occupied by a gas is inversely proportional to the pressure (P). (PV = constant.)

(2) At constant pressure, the volume of a gas is proportional to the absolute temperature (0° C. is 273° on the absolute temperature scale).

(3) The partial pressure or tension of a gas, in a mixture of gases having no action on one another, is equal to that which the particular gas would exert did it alone take up the space occupied by the mixture; in other words, in a mixture of gases at a certain pressure, the total pressure is divided between the different gases in proportion to their relative volumes.

[1] This is shown by the following example:

Volume of tidal air = 500 c.c. CO_2 content per cent. of inspired air = 0; of expired air = 4·0; of alveolar air = 6·0. Let x = vol. of dead space.

Amount of alveolar air in c.c. present in expired air is $500 - x$.

Volume of CO_2 in c.c. in expired air = vol. of CO_2 in dead space + vol. of CO_2 in exhaled alveolar air.

Substituting:

$$\frac{500 \times 4}{100} = x \times 0 + \frac{(500 - x) \times 6}{100}$$

$$2000 = 3000 - 6x.$$

$$6x = 1000 \text{ c.c.}$$

$$x = 166 \text{ c.c.}$$

[2] Crockford and Knight, *Physical Chemistry*. N.Y., 1950.

(4) The amount of gas going into solution in a liquid which has no chemical attraction for the gas depends on its *solubility* and is proportional to the partial pressure of the gas.

The solubility of the respiratory gases in c.c. per 100 c.c. of water, plasma, and blood at *38° C.* and *760 mm. Hg* pressure is as follows :

	WATER.	PLASMA.	BLOOD.
Oxygen .	2·37	2·30	2·30
Nitrogen .	1·20	1·20	1·10
CO_2 . .	55·5	54·1	54·1

Note that the solubility is diminished by the presence of salts in plasma and blood, and that CO_2 is far more soluble than either O_2 or N_2.

Tension of a Gas in a Liquid.—The meaning of this expression is a little difficult to grasp and is best illustrated as follows : if a liquid is exposed for a sufficient length of time to a certain pressure (tension) of a gas, *e.g* if water is exposed to a pressure of CO_2 of 40 mm. Hg, the molecules of the gas (CO_2) pass into solution in the liquid until equilibrium is reached; the tension of the gas (CO_2) dissolved in the water is then the same as the tension of the gas outside, *i.e.* it is 40 mm. Hg. One must picture the dissolved molecules of the gas as having been " bottled " in the liquid under pressure, and that these molecules are " trying " to escape from the liquid with a pressure equal to that with which they were driven in. If the liquid is exposed to a higher CO_2 pressure, *e.g.* 46 mm. Hg, more CO_2 passes into solution ; if to a lower pressure, *e.g.* 35 mm. Hg, CO_2 leaves the solution ; if to the same pressure (40 mm. Hg) no loss or gain of CO_2 takes place. Two gases may be dissolved in a liquid at the *same tension* and yet the *volumes* in solution may be very different, depending on the respective solubility of the gases. Thus (at 38° C.) 100 c.c. water which are exposed to a CO_2 pressure of 40 mm. hold about 2·5 c.c. CO_2 in solution ; but exposed to the same pressure of O_2, only about 0·12 c.c. of oxygen passes into solution (though the tension of the two gases in the water is exactly the same). The tension of a gas in a fluid may be determined by exposing it to various gas mixtures of different compositions, and finding with which it is in equilibrium, *i.e.* when no gas is given off or taken up. Krogh's method (p. 367) may be employed in experiments on the blood of animals.

Gaseous Interchanges in the Lungs.—The gaseous interchange between the alveoli (air-sacs) and the pulmonary capillary blood depends entirely on the physical process of *diffusion*, *i.e.* on the difference of gas pressures in the blood and the alveolar air. There is no active intervention (so-called gas secretion) by the pulmonary epithelium. The velocity of gaseous exchange between the lungs and the blood depends on : (i) *pressure* difference, (ii) *solubility* of the gas in blood, (iii) the properties of the *membrane*.

$$\text{Velocity of exchange} = \frac{\text{Pressure gradient} \times \text{solubility of gas}}{\sqrt{\text{density of gas}}} \times K$$

$$\text{where } K = \frac{0{\cdot}139 \times \text{area of lung surface}}{\text{thickness of lung membrane}}.$$

The pressure gradient in the case of CO_2 is one-tenth that of O_2. the density of CO_2 is 44 compared with 32 for O_2, but CO_2 is about twenty-five times as soluble as O_2. CO_2, therefore, diffuses out into the alveoli more quickly than O_2 diffuses into the blood in spite of the disparity in the pressure gradients for the two gases.

The area of the lung surface is assessed at 55–70 sq. metres; the surface area of the pulmonary capillaries is said to be 40 sq. metres; the volume of blood present at any moment in the capillaries is about 60 c.c. An extremely fine stream of blood presenting an enormous surface and flowing through very short capillaries in less than 1 second (p. 307) is thus exposed to the alveolar air and is separated from it by two very thin membranes, *i.e.* the pulmonary epithelium and the capillary endothelium.

DETERMINATION OF GAS PRESSURES IN ARTERIAL AND MIXED VENOUS BLOOD.—By means of *Krogh's aerotonometer* the tension of the gases in arterial and venous blood can be determined in experimental *animals*. A cannula is introduced into the pulmonary artery or into a systemic vessel. The blood to be examined is then allowed to play for many minutes on a small air bubble in the apparatus until equilibrium is established. As the volume of the gas bubble is small and that of the blood which comes in contact with it is very large, the gases in the bubble come into equilibrium with the gases in the blood. The bubble is analysed, the gas tensions in it are calculated, and thus the blood-gas tensions are determined.

In *man* the gas tensions in the *mixed venous blood*, *i.e.* that in the right heart or pulmonary artery, may be determined directly by collecting a sample of blood from the right auricle (p. 279). *Arterial* blood may be taken directly from a peripheral artery like the radial into a special syringe; a gas bubble is then introduced into the blood in the apparatus and allowed to remain in contact with it until equilibrium is established; the bubble is then analysed.

The following Table shows the average results which are obtained:

	O_2 Tension.	CO_2 Tension.
Alveolar air	100 mm. Hg	40 mm. Hg
Mixed venous blood (at rest)	40 mm. ,,	46 mm. ,,
Arterial blood	100 mm. ,,	40 mm. ,,

The differences in gas tension between alveolar air and blood are such as enable gaseous interchange to take place by diffusion, *i.e.* difference of gas tension on the two sides of the pulmonary epithelium. The oxygen tension in alveolar air is about 100 mm. Hg, in venous blood it is 40 mm. Hg, *i.e.* there is a difference of pressure of over 60 mm. Hg on the two sides of the membrane. The corresponding CO_2 tensions are 40 mm. and 46 mm.; the difference of pressure in this case is quite small, only 6 mm. Hg. As CO_2 passes very readily through the pulmonary epithelium, it can be calculated that this small difference of tension is sufficient to eliminate from the body about 10 litres of CO_2 per minute. A difference of CO_2 tension across the lung membrane of 0·12 mm. Hg is sufficient to enable the resting output of CO_2 (200–250 c.c. per minute) to take place.

Intrapleural Pressure.[1]—The intrapleural space is very small in the living animal, as the visceral and parietal layers of the pleura are practically in contact. The pressure in the pleural cavity is *subatmospheric*; it is −5 mm. Hg (*i.e.* 5 mm. *less than atmospheric*) during expiration and −10 mm. Hg during inspiration. The pressure in the lungs is about 1 atmosphere, as they are in free communication with the outside air. The full intrapulmonary pressure is not transmitted to the pleural cavity, however, owing to the *elastic recoil of the lungs* (Fig. 219). The more the lungs are stretched, the greater is their tendency to recoil. Consequently, the more the lungs are distended by a deep inspiration, the more subatmospheric (*i.e.* more negative) does the pressure in the pleural cavity become. As previously pointed out (p. 275), the negative intrapleural pressure aids the venous return to the chest. If the chest wall is *widely* opened, air enters the pleural cavity and the lungs collapse (*vide infra*).

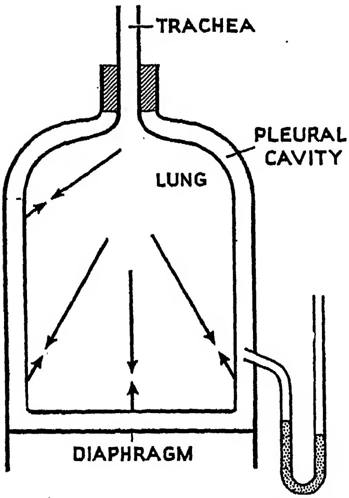


FIG. 219.—Mode of Production of Negative Intrapleural Pressure.

Pneumothorax.[2]—By pneumothorax is meant the presence of air (or other gas) in the pleural cavity. A pneumothorax may be (i) *open*, when air can pass freely in and out of the pleural cavity through an opening in the chest wall or in the lung itself; (ii) *closed*, when the air is imprisoned in the pleural cavity; (iii) *valve type*, when air enters freely (during inspiration) but cannot pass out (during expiration).

EXPERIMENTAL OPEN PNEUMOTHORAX.—In some animal species, *e.g. dog*, the mediastinal structures possess negligible resistance—they are *not* rigid and unyielding; the pressures in both pleural cavities are consequently always the same within practical limits. Thus, in the dog, when a hole is made in the chest wall, air enters the pleural cavity on that side, the increased pressure is readily transmitted to the opposite side too, and there is an equal rise of pressure on the two sides. The response to this hindrance to breathing consists of an increase in the depth of respiration. These points are well illustrated in Fig. 220. An opening in the left pleura was made at the point indicated by the first arrow. Immediately after, there is a change of pressure in *both* pleural cavities, from one which is wholly negative (during both phases of respiration) to one which is mostly positive. Respiration is increased in amplitude and slower; in spite of the larger chest movements very little air enters the lungs, and the tracheal pressure is practically steady at the atmospheric pressure line. On closing the opening (second arrow) there is an immediate response in both pleural cavities with restoration to a large extent of the negative pressure. More air now enters the lungs, as is

[1] Prinzmetal and Kountz, *Medicine*, 1935, *14*, 457

[2] Graham and Bell, *Amer. J. med. Sci.*, 1918, *156*, 850. Christie and McIntosh, *Quart. J. Med.*, 1936, *5*, 445. Alexander, *Collapse Therapy*, London, 1937

shown by the larger tracheal pressure oscillations. Breathing still remains slow.

It is obvious that under the experimental conditions described, when the thorax is enlarged during inspiration, air will enter by *both available inlets*, *i.e.* the glottis and the new opening in the chest wall. It must be remembered that the elasticity of the lung and the frictional resistance in the air passages always impede the entry of air into the lungs; if the artificial opening is big enough, practically no air enters the lungs even when the breathing becomes maximal in depth; instead, all the air enters by the pneumothorax opening, and death from asphyxia rapidly supervenes. The smaller the new opening and the greater the compensatory increase in breathing, the more likely is air to enter the lungs. Graham has calculated (*on the unproven assumption that conditions in man are the same as in the dog*) that in a person who could increase his tidal air to a vital capacity of say 3500 c.c., the largest size of opening in the chest compatible with life is 5×10 cm. (area of glottis is 2·25 square cm.); in other words, just enough air could be introduced into the lungs under these circumstances to ward off asphyxia, but only for a short time (cf. p. 370).

Open pneumothorax involves further complications: (i) constant risk of infection of the exposed lung; (ii) rapid heat loss—the temperature of an animal with an open pneumothorax may fall by 3·5° F. in 45 minutes, while if the abdomen is opened and the intestines are drawn out, the temperature only falls by 1° F. in the same period; (iii) as the mediastinal structures flap from one side

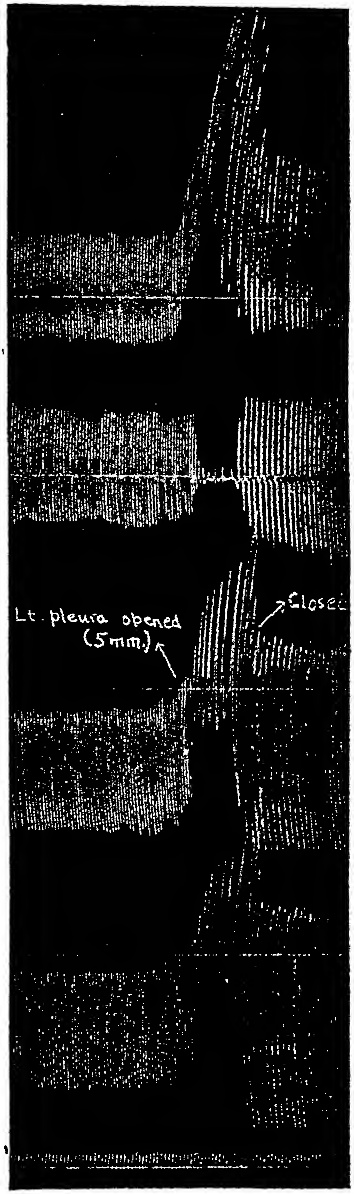


FIG. 220.—Acute Experimental Open Pneumothorax. (After Graham and Bell, *Amer. J. med. Sci.* 1918.)

Dog. Ether Anæsthesia. Records from above downwards are respiratory movements, tracheal pressure, left intrapleural pressure, right intrapleural pressure. The horizontal line on each record represents the atmospheric pressure level. Time at bottom in seconds.

to the other with inspiration and expiration, the heart's action may be impeded.

When an open pneumothorax is converted into a closed one, immediate and striking relief is obtained (Fig. 220). This is not due to absorption of the air (though this does occur gradually into the blood), because the improvement is so very rapid in onset. The explanation is that now, when the chest is enlarged during inspiration, the additional space made available is taken up by air which *enters the lungs only* ; asphyxia now occurs only when so much chest space is taken up by the pneumothorax that the minimum air requirement cannot enter the lungs.

PNEUMOTHORAX IN MAN.[1]—It is believed that the *normal* mediastinal structures in man (as in dog) possess little resistance, with the result that the pressure in both pleural cavities tends to be approximately the same. But if the mediastinum is diseased it becomes much more resistant and unyielding; pressure changes are then not necessarily transmitted readily from one pleural cavity to the other. There appear to be no observations in man as to what happens to the pleural pressure on the intact side in cases of traumatic open pneumothorax. The pressure in the open pleural cavity is of course atmospheric ; in the intact cavity it is probably subatmospheric to a degree which depends on the state of the mediastinum. Under these conditions the chest cavity may be regarded as divided into two distinct compartments which function largely independently one of the other. As the pressure in the pleural cavity and that inside the lung on the *affected* side are both atmospheric, the lung on that side *collapses* by virtue of its own elasticity. During inspiration, the enlargement of the intact half of the thoracic cavity leads to normal expansion of the lung on that side; the enlargement of the chest on the affected side leads to air being aspirated into the open pleural cavity from the exterior, the lung on that side remaining collapsed, no air entering it from the trachea. During expiration some of the air expelled (at slightly positive pressure) from the normal side may pass into, and slightly expand, the collapsed lung, this air being sucked out again during inspiration. The mediastinal structures swing from side to side with the phases of respiration ; as a result there is periodic kinking of the thin-walled great veins, resulting in interference with the venous return and decrease in the cardiac output. Compensation is effected by roughly doubling the amplitude of the movements of the chest wall and so doubling the ventilation of the normally functioning lung which makes up for the loss of the affected lung. The increased pulmonary ventilation is partly brought about by the combined action of anoxia and CO_2 accumulation ; in addition afferent impulses pass up from the collapsed lung along the vagus nerve reflexly stimulating the respiratory centre (p. 457). If the circulation through the collapsed lung is blocked and the whole right ventricular output passes through the ventilated lung then arterialization of the blood may be normal so long as the effective pulmonary ventilation is adequate ; to the extent, however, that any circulation continues through the collapsed lung, the venous blood is shunted unchanged from the right to the left heart leading to anoxic anoxia and CO_2 retention, which intensify the respiratory distress (cf. p. 448). Closure of the opening in the chest wall relieves the symptoms for the reasons explained above.

[1] Sauerbruch and O'Shaughnessy, *Thoracic Surgery*, London, 1937.

In *valvular* pneumothorax the pressure on the affected side may rise as high as +20 mm. Hg. Under these conditions the mediastinum is markedly *displaced* and presumably the pressure in the intact pleural cavity is raised; the capacity of the chest cavity on that side is decreased. An even greater effort is needed to ventilate the intact lung adequately for bodily needs. The raised intrapleural pressure combined with the mediastinal displacement probably reduces the blood flow in the great veins and the filling of the auricles, thus leading to a diminished pulmonary blood flow and a decreased systemic cardiac output. Because of these additional complications great respiratory distress is experienced; there is marked cyanosis and the visible veins are greatly engorged. Marked relief is experienced when the imprisoned air is withdrawn to lower the intrapleural pressure.

In *closed* pneumothorax such symptoms as may be present are due mainly to the collapse of the affected lung (p. 448). But during inspiration the enlargement of the chest cavity leads to an equivalent volume of air being drawn into the lungs so that adequate arterialization of the blood may often be effected.

Absorption of Gas from Pneumothorax.—Gases are frequently introduced into the pleural cavity to cause partial collapse of the lungs as a therapeutic measure in certain forms of pulmonary tuberculosis; the condition produced is called an *artificial pneumothorax.* As there is no opening in the chest wall, large volumes of gas can be thus introduced without causing any distress (p. 370); the gas (usually air) is always slowly and finally completely absorbed into the blood. The visceral layer of the pleura is permeable to the respiratory gases (O_2, CO_2, and N_2); gaseous interchange occurs between the gas in the pleura and the venous blood in the pulmonary capillaries to establish pressure equilibrium on the two sides of the membrane. The pressure of the pneumothorax gases is about atmospheric (760 mm. Hg); the total pressure of all the gases in *arterial* blood is also atmospheric as it is in pressure equilibrium with the gases in the alveolar air; the *total gas pressure in the venous blood is, however, substantially less than atmospheric* (and therefore less than that in the pneumothorax) because it has suffered a large fall of oxygen pressure, but has only received a small increase in CO_2 pressure. This is shown in the following table (pressures in mm. Hg):

Pressures.	Arterial Blood in mm. Hg.	Venous Blood in mm. Hg.
Total	760	706
Water vapour	47	47
Dry gases	713	669
Oxygen	**100**	**40**
Carbon dioxide	**40**	**46**
Nitrogen	573	573

As a result, gaseous diffusion takes place from the pneumothorax into the pulmonary venous blood until complete absorption occurs.

In practice, when an artificial pneumothorax is produced therapeutically, about 400 c.c. air are injected on the first occasion; second and third "refills" of the same amount are made after two and three days respectively. Later,

larger amounts (about 700 c.c.) are injected at longer intervals (every two or three weeks). If the visceral pleura is thickened as a result of inflammatory change it becomes impermeable to gases, and the pneumothorax cannot be absorbed. If the lung is fibrosed and cannot expand, only a small initial absorption of gas occurs, as the lungs cannot enlarge to fill the space provided and thus restore the pleural pressure to its original value.

Determination of Gaseous Metabolism.—It is possible by means of simple respiratory methods to determine the *oxygen consumption*, the *CO_2 output*, the *respiratory quotient*, and the *metabolic rate* under various conditions of health and disease. A mouthpiece with two suitably arranged valves is employed; atmospheric air is breathed in through one, and the expired air is breathed out through the other. A large canvas Douglas bag of 100 litres capacity or more is connected to the expiratory side of the mouthpiece, and expired air is collected for a suitable time, usually 5 minutes. The volume of the expired air is found by passing it through a special gas meter (spirometer), and its composition is determined by analysing a sample in the Haldane gas analysis apparatus. This apparatus consists of a 10 c.c. graduated burette (into which the gas sample is introduced) connected to two bulbs—one containing potash for the absorption of CO_2 and the other alkaline pyrogallol to absorb O_2. The initial volume of the sample is read off on the burette (*e.g.* 10 c.c.) and the decrease in volume observed which results from the successive absorption of CO_2 and O_2. The percentage amount of these gases in the sample is thus determined. [The Haldane gas apparatus is used for the analysis of any suitable gas mixture—*e.g.* inspired, expired, or alveolar air.] From the data thus obtained the CO_2 output and O_2 consumption per minute may be determined thus:

Pulmonary ventilation per minute (from spirometer readings reduced to N.T.P) = 5 litres.

Inspired air: $O_2 = 21$, $CO_2 = 0$ per 100 c.c.

Expired air (from analysis of sample): $O_2 = 16$, $CO_2 = 4$ per 100 c.c.

For every 100 c.c. air respired CO_2 given out is $4 - 0 = 4$ c.c.; O_2 absorbed is $21 - 16 = 5$ c.c. As the minute ventilation is 5000 c.c.

$\therefore$ O_2 consumed is 250 c.c.; CO_2 output is 200 c.c. per minute.[1]

The *O_2 consumption* can be determined directly by means of the Benedict-Roth apparatus (p. 375).

Respiratory Quotient.[2]—The *respiratory quotient* (R.Q.) is the *ratio of the volume of CO_2 evolved from the lungs/the volume of O_2 absorbed from the lungs in a given time.*

In the above example it is $\frac{200}{250} = 0{\cdot}8$.

The respiratory exchanges and the R.Q. can thus be determined under varying conditions of diet, health, and disease. Two additional points should be stressed.

(i) It is difficult to determine the R.Q. accurately in untrained subjects because they almost invariably overventilate as soon as they begin to breathe

[1] For full details of calculation which are given in a simplified and not entirely exact form in the text, see Douglas and Priestley, *Human Physiology*, 3rd edn, Oxford, 1948.

[2] Richardson, *Physiol. Rev.*, 1929, *9*, 61.

through valves; they thus eliminate more CO_2 than normally and so yield a misleading high R.Q. value. It is desirable that a preliminary practice period of 15 minutes breathing through the valves should precede the actual determination.

(ii) The R.Q. is only a *ratio*: if the O_2 consumption or CO_2 output are equally increased or decreased the R.Q. is unaltered. Thus if the ventilation is increased because of an increase in metabolism (*e.g.* in exophthalmic goitre) the R.Q. is not necessarily altered, because both O_2 consumption and CO_2 output rise. The R.Q. would change only if there was some other associated change, *e.g.* in the food mixture used by the body, or in the blood reaction.

The respiratory quotient has been intensively studied in many conditions in health and disease and the main results are summarized below; but it must be confessed that the respiratory quotient frequently gives no information about metabolic processes in the body; on the contrary, a *knowledge of the metabolic processes taking place is generally necessary to interpret the R.Q.* Furthermore the R.Q. throws no light on the stages of intermediate metabolism.

(1) Effect of Combustion of Foodstuffs, Intermediate Metabolism and Diet.—If pure foodstuffs are burnt in a bomb calorimeter their respiratory quotient can be measured directly.

(i) With pure carbohydrate it is 1.

$$C_6H_{12}O_6+6O_2=6CO_2+6H_2O$$

i.e. the volume of CO_2 evolved is equal to the volume of oxygen used.

(ii) With the fatty acids derived from food fat the R.Q. is about 0·7.

$$C_{17}H_{35}COOH+26O_2=18CO_2+18H_2O.$$
$$R.Q.=18/26=0{\cdot}693.$$

Fat contains very little oxygen which must consequently be provided in sufficient amount to oxidize both the hydrogen and the carbon of the fat molecule.

(iii) With pure protein it is about 0·8.

It may be safely assumed that when fat or carbohydrate is *completely* combusted in the body the R.Q. is still 0·7 or 1·0 respectively, *whatever the intermediate stages through which the foodstuff passes.* This fact is illustrated by the following example. Let us suppose that food fat or reserve fat is converted in the body to carbohydrate which is then combusted. The reactions involved in these transformations can be crudely summarized as follows:

(*a*) $\underset{\text{stearic acid.}}{C_{18}H_{36}O_2}+8O_2=\underset{\text{glucose.}}{3C_6H_{12}O_6}.$

(*b*) $3C_6H_{12}O_6+18O_2=18CO_2+18H_2O.$

Adding the two reactions to get the R.Q. of the whole process:

(*c*) $C_{18}H_{36}O_2+26O_2=18CO_2+18H_2O,$

i.e. R.Q.$=\frac{18}{26}=0{\cdot}693$, which is the R.Q. of the direct complete combustion of fat.

When carbohydrate is converted into fat which is *stored* and not combusted the reaction can be summarized as follows:

(*d*) $3C_6H_{12}O_6=C_{18}H_{36}O_2+8O_2.$

O_2 is thus released which is used by the cells for oxidation processes, so reducing the O_2 intake from the lungs without decreasing tissue oxygen consumption or CO_2 formation or elimination. The R.Q. is the ratio of CO_2 output from the lungs/O_2 intake from the lungs; as the denominator (the O_2 intake *in the lungs*) has fallen owing to reaction (*d*) the R.Q. value for the *whole* metabolism of the body is raised.

Conversely reaction (*a*) *supra* while taking place alone involves additional uptake of O_2 in the lungs over and above that needed for oxidation of the foodstuffs; in the ratio CO_2/O_2 the denominator is increased and so the R.Q. of the *whole* metabolism of the body is lowered.

It follows therefore that if the R.Q. is determined from measurements extending over *long continuous periods*, the value obtained will approximately indicate the mixture of foodstuffs undergoing oxidation, *e.g.* if carbohydrate predominantly is being oxidised the R.Q. will approximate to 1, if fat the R.Q. will be about 0·7. There is a tendency also for the tissues normally to utilise preferentially the foodstuff most freely available, so the R.Q. will tend to vary with the composition of the food. The non-nitrogenous residues of protein enter the common metabolic pool and are then treated essentially as intermediates from the oxidation of carbohydrate and fat.

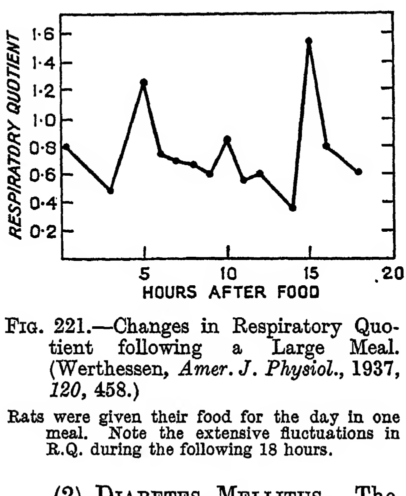


FIG. 221.—Changes in Respiratory Quotient following a Large Meal. (Werthessen, *Amer. J. Physiol.*, 1937, *120*, 458.)

Rats were given their food for the day in one meal. Note the extensive fluctuations in R.Q. during the following 18 hours.

The R.Q. determined over a *short* period will depend on the R.Q. of the *intermediate processes* through which all the foodstuffs are passing during the time. Thus rats were given the whole of their daily diet in one meal; the R.Q. subsequently determined at intervals varied between 0·27 and 1·6 (Fig. 221).

(2) DIABETES MELLITUS.—The interpretation of the R.Q. in diabetes mellitus is considered on p. 920.

(3) ALTERATIONS IN PULMONARY VENTILATION.—The R.Q. is affected (independently of any change in the nature of foodstuffs oxidised) when, in conditions of stable metabolism, the *pulmonary ventilation is altered either voluntarily or secondarily to a rise of body temperature or to fluctuations in the H^+ ion concentration of the blood* (cf. Fig. 57). Thus:

(i) Voluntary hyperpnœa washes out excessive quantities of CO_2 (without increase in O_2 consumption); the R.Q. may rise considerably above unity.

(ii) In acidæmia from any cause, *e.g.* ingestion of NH_4Cl, or secretion of the alkaline intestinal juices, there is hyperpnœa (p. 396) with increase in CO_2 output but again without corresponding rise in O_2 consumption; the R.Q. therefore rises. The same applies to the hyperpnœa of raised body temperature (p. 478).

(iii) Conversely in alkalæmia, *e.g.* from ingestion of $NaHCO_3$, the breathing is depressed, CO_2 is retained in the body, and the R.Q. falls (p. 397).

(4) VIOLENT EXERCISE.—During very violent exercise, lactic acid enters

the blood stream and breaks up the plasma $NaHCO_3$, liberating large additional volumes of CO_2 which are eliminated from the lungs; the R.Q. may then exceed 2. During recovery from exercise, CO_2, which is derived from oxidative processes in the muscles, is retained in the blood in large amounts to reform the bicarbonate, and the R.Q. falls to a very low value (p. 439). The R.Q. is not affected in a constant manner by *moderate* exercise because there is no acidæmia and roughly the resting mixture of foodstuffs is being metabolized (cf. p. 438).

Metabolic Rate.—From a knowledge of the respiratory exchanges, the *metabolic rate, i.e.* the energy produced in the body, can be indirectly determined. It may be assumed for the sake of simplicity that all the energy is derived from varying proportions of carbohydrate and fat. The metabolism of protein, apart from the nitrogenous fraction which is excreted, is essentially that of the carbohydrate and fatty acids (or their intermediates) derived from the residues.

If different proportions of carbohydrate and fat are burnt in a calorimeter with 1 litre of oxygen, it is found that pure carbohydrate (R.Q.=1) yields about 5 Calories,[1] pure fat (R.Q.=0·7) yields about 4·8 Calories, and various mixtures of carbohydrate and fat (R.Q. between 0·7 and 1) give between 4·8 and 5 Calories as shown in the Table below.

R.Q.	Calories evolved per Litre of O_2 consumed.
0·71	4·795
0·75	4·829
0·80	4·875
0·85	4·921
0·9	4·967
0·95	5·012
1·0	5·058

It is assumed, probably unjustifiably, that the R.Q. values obtained in man have a similar significance. The R.Q. being known, the heat value of 1 litre of oxygen is deduced from the table; if the *oxygen consumption* is also known the heat production can be determined. Thus, if the R.Q.=0·8 (calorific value per 1 litre O_2=4·875), and the O_2 consumption per minute is 0·25 litre, then the heat production per minute is $0{\cdot}25 \times 4{\cdot}875 = 1{\cdot}22$ Calories.

Clinically the metabolic rate under conditions of complete rest and fasting (*basal metabolic rate,* p. 377) is calculated from the *oxygen consumption alone.* This is determined most readily by means of the Benedict-Roth apparatus which consists essentially of an accurately constructed tank filled with oxygen and suspended in water. It is connected by means of tubing through a soda-lime tower to the patient's mouth (the nose is clipped). The

[1] One calorie is the amount of heat required to raise the temperature of 1 g. of water by 1° C.; a more precise term is 1 g.-calorie. It may be abbreviated to 1 cal. [In physiological texts it is sometimes referred to as 1 small calorie.]

One Calorie=1000 calories. It may be abbreviated to 1 Cal. A more precise term is 1 kg.-calorie. [It is sometimes referred to as 1 large calorie.]

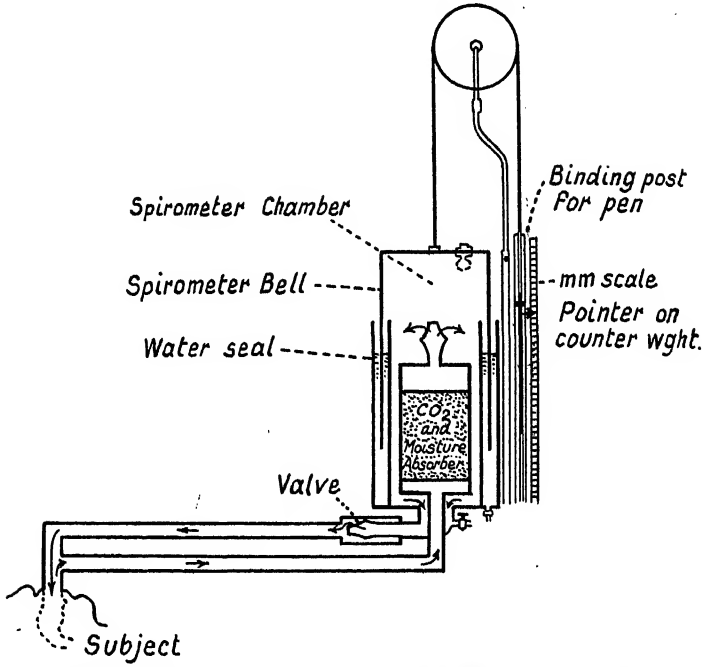


Fig. 222.—Benedict-Roth Apparatus for determining the Metabolic Rate. (Beaumont and Dodds, *Recent Advances in Medicine*, Churchill.)

The arrows in the diagram show the direction of the flow of air during respiration.

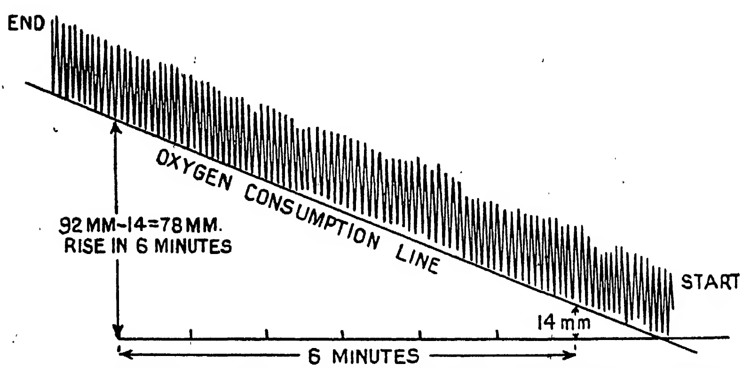


Fig. 223.—Graphic Record of Determination of Metabolic Rate. (Beaumont and Dodds, *Recent Advances in Medicine*, Churchill.)

The record (which reads from *right to left*) is obtained using the apparatus shown in the preceding figure This particular instrument is so constructed that each 1 mm. fall of the spirometer bell (and corresponding rise in the record) represents 20·73 c.c. oxygen used (this value is not corrected to N.T.P.). If the calorific value of oxygen is assumed to be 4·82 Cal. per litre, then 20·73 c.c. of oxygen represent the evolution of 0·1 Cal. In this experiment the rise in the record in 6 minutes is 78 mm. corresponding to an uncorrected heat output of 78 × 0·1 Cal. = 7·8 Cal., or 1·3 Cal. per minute.

patient rebreathes from the tank; the CO_2 formed is removed by means of the soda lime, and the decrease in the volume of oxygen in the tank is a direct measure of the oxygen consumption. The CO_2 output is *not* determined, and the respiratory quotient is assumed to be 0·75 (as the subject is fasting), corresponding to an output of 4·82 Calories per litre of oxygen consumed (Figs. 222 and 223).

Basal Metabolic Rate (Basal Metabolism.)[1]—By this is meant the energy output of an individual under *standardized resting conditions, i.e.* at complete bodily and psychical rest, 12–18 hours after a meal (post-absorptive period) and in an equable environmental temperature. It can be determined by the methods just described. Under such conditions a proportion

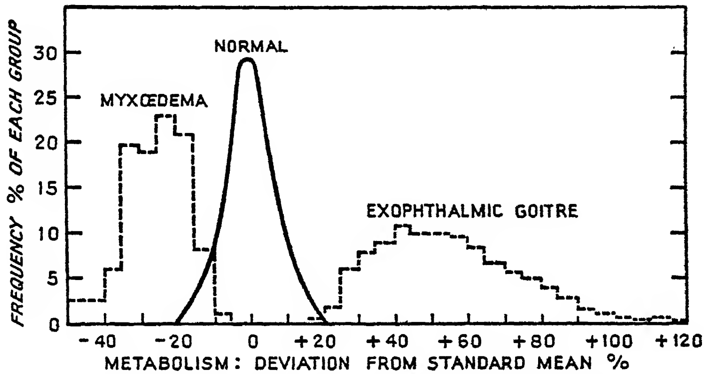


FIG. 224.—Basal Metabolism in Myxœdema and Exophthalmic Goitre compared with group of Normal Individuals. (Du Bois, *Ann. int. Med.*, 1938, *12.*)

Ordinate: percentage frequency for each group.
Abscissa: metabolic rate as percentage deviation from accepted normal average standard (=0).

Note that almost all the cases of myxœdema fall below the normal range; those of exophthalmic go (Graves' disease) are all above the normal range.

of the energy liberated is used to maintain the activities of vital organs like the heart, brain, or glands, but the greater part is converted into heat so as to maintain body temperature and prevent it from falling below the normal level. The basal metabolism is *not* the lowest possible level to which the metabolism can fall in the subject; it may be lowered further, for example, by starvation or thyroid deficiency; it is simply as stated the metabolism under standard resting conditions.

Clinically the basal metabolic rate (B.M.R.) is expressed as a percentage above or below the theoretical normal standard for the individual, taking into account his age, height, weight, etc. (p. 378). Thus a B.M.R. of +50 means one which is 50% above the normal average for that person. Fig. 224 shows that the *normal* range of metabolism in people of identical physical status may vary commonly by ± 10% and rarely even by as much as ± 20%.

[1] Boothby and Sandiford, *Physiol. Rev.*, 1924, *4*, 69. Du Bois, *Basal Metabolism in Health and Disease*, 3rd edn., Phila., 1936. Peters and Van Slyke, *Quantitative Clinical Chemistry*, 2nd edn., vol. I, pt. i, p. 1, 1946.

Factors influencing Metabolic Rate.—(1) SURFACE AREA.—The basal metabolism is most closely related to the *surface area* and is less directly related to *height* or *weight*. The surface area can be calculated from the following formula if the height and weight are known (Du Bois): $A = W^{0.425} \times H^{0.725} \times 71{\cdot}84$, where A = surface area in sq. cm., W = weight in kg., and H = height in cm.

In the *male* 40 Calories, and in the *female* adult about 37 Calories, are given off every hour per square metre of body surface (the surface area of an average adult is about 1·8 sq. metres); or, expressed in terms of body weight, the basal metabolism amounts to about 1 Calorie per kg. per hour. The values quoted for metabolic rate are *average* normal values; an average implies that higher and lower values occur. The range of normal variation was referred to above.

(2) AGE.—The basal metabolism is considerably greater per sq. metre of surface in *children* than in adults; there is a further gradual fall in the metabolism during adult life as age advances. These facts are well shown in the Table below.

Age in years.	B.M.R. in Calories per sq. metre per hour.	
	Male	Female
2	57·0	52·5
6	53·0	50·6
8	51·8	47·0
10	48·5	45·9
16	45·7	38·8
20	41·4	36·1
30	39·3	35·7
40	38·0	35·7
50	36·7	34·0
60	35·5	32·6

Fig. 225 shows two B.M.R. determinations carried out by Magnus Levy (the pioneer investigator in this field) on himself at an interval of 50 years; the decrease in his B.M.R. (per hour, per kg. per hour and per sq. metre per hour) is strikingly demonstrated.

(3) STARVATION or *prolonged undernutrition* damps down the metabolic rate. For example, in a man who had fasted 31 days, the daily basal metabolism diminished from 958 to 737 Calories per sq. metre of surface area—a fall of over 20% (p. 1048). In poorly nourished patients the reduced body weight can finally be maintained on considerably less than the standard basal caloric requirements.

(4) BODY TEMPERATURE.—For every rise of 1° F. in the internal temperature of the body, the basal metabolism increases by 7%. The chemical reactions of the body, like those occurring in a test-tube, are speeded up by a rise of temperature. Thus, a patient suffering from pneumonia with a temperature of 105° F. (about 7° F. above the normal) would have an increase of 50% in his metabolism (and in his pulmonary ventilation) because of the fever alone.

(5) External Temperature.—Exposure to *cold* increases the metabolism; there is consequently increased heat production which helps to maintain the normal body temperature (p. 480). Exposure to external heat of brief duration has little effect on metabolism, as compensation is effected mainly by increasing heat loss; if the exposure is prolonged, a gradual fall in the metabolic rate takes place (p. 476).

YEAR	AGE YR.	OXYGEN c.cm. per minute	CARBON DIOXIDE c.cm. per min.	RESPIRATORY QUOTIENT	HEIGHT Cm.	WEIGHT Kg.	SURFACE AREA Sq. m.	CALORIES per hr.	CALORIES Kg./hr.	CALORIES Sq.m. per hr.
1891	26	231	192	0·83	167·0	67·5	1·76	67	0·99	38·1
1941	76	176	158	0·90	165·5	60·0	1·65	52	0·87	31·5
DIFFERENCE	50	−24%				−11%	−6%	−22%	−12%	−17%

Fig. 225.—A Likeness of Professor Magnus Levy and the Changes in his Basal Metabolic Rate after 50 years.

(6) Ductless Glands.—(i) The active principle of the thyroid gland—*thyroxine*—acts as a general catalyst, speeding up the metabolic activities of the tissues (p. 977).

Thus in exophthalmic goitre, in which there is increased secretion of the active principle of the gland, the basal metabolism may increase in a severe case up to double the normal (*i.e.* the B.M.R. is up to +100 (Fig. 224)). In myxœdema, in which there is lessened secretion of thyroxine, metabolic

activity may be depressed to 60% or 70% of the normal (*i.e.* the B.M.R. is —30 or —40; the last value is about the lowest ever observed clinically and represents the *minimum* metabolic level (Fig. 224).

(ii) *Adrenaline* increases the metabolic rate, but to a less extent than thyroxine. The injection of 1 mg. adrenaline in man increases heat production up to about 20% for a few hours only.

(iii) The *anterior pituitary* influences the metabolic rate indirectly through its thyrotrophic hormone.

(7) In certain other conditions there is an increased metabolic rate for no very clear reason. This is seen in both splenomedullary and lymphatic *leukæmia*, in which the increase may be of the same order as that found in Graves' disease. A less marked increase is found in *pernicious anæmia* and in *congestive heart failure*.

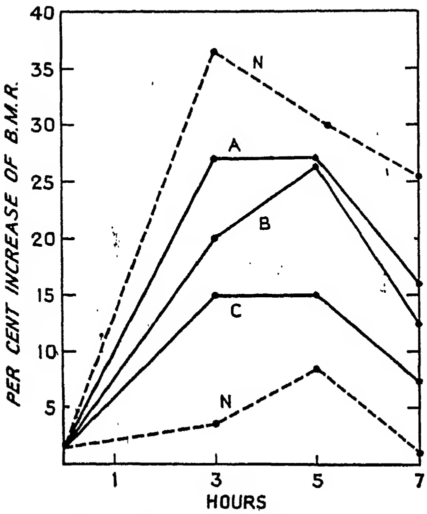


FIG. 226.—Effect of Ingestion of Protein on Metabolic Rate (Specific Dynamic Action) in Man. (Modified from Abel, *Amer. J. med. Sci.*, 1943, *205*, 417.

The subjects ate a meal containing 125 g. of protein and 5 g. of fat. Ordinates show percentage increase in metabolic rate. The upper and lower dotted lines (N, N) show the upper and lower limits of normal responses. Curves A, B, C show the responses in three normal subjects.

(8) The taking of food stimulates metabolism. This effect is not equally marked with all classes of foodstuffs, being least with carbohydrate and fat, and greatest with protein. If 125 g. of protein in the form of meat are eaten at one meal, the metabolism rises to reach a maximum after 3–5 hours and then slowly declines (Fig. 226); the *peak* percentage increase is 10–35%; the *total* increase in metabolism may average 20% over a period of 4–6 hours. If carbohydrate and fat are eaten in amounts of equal calorific value, the metabolism increases by only 5–10%. The more striking effect of protein is termed its *specific dynamic action*. The increase in metabolism is attributed to a stimulating action on cellular metabolism of the fatty acid residues which are left after the NH_2 groupings have been removed from the amino-acids. The stimulating action of a carbohydrate meal is attributed to extra carbohydrate being burnt to provide the energy necessary to convert some of the glucose into a glycogen store; this energy is really being *stored* and not expended because the resulting glycogen has a higher chemical energy level than the glucose from which it is derived. If an ordinary mixed diet is taken, the metabolism is increased by 50–150 Calories daily.

(9) Lastly, and most important, there is an increase in the metabolism with muscular work. During very violent exercise the oxygen consumption per minute may rise from 250 c.c. to 4 litres, or even more, *i.e.* the metabolism may increase over 16 times (p. 436).

Artificial Respiration.[1] — The methods commonly employed in man will first be described; their uses and comparative advantages will then be considered.

SCHÄFER'S METHOD.—"The method consists in laying the subject in the *prone* posture, preferably on the ground, with a thick folded garment underneath the chest and epigastrium." The head is turned to one side. "The operator puts himself in a position athwart or at the side of the subject, facing his head and kneeling upon one or both knees, and places his hands on each side over the lower part of the back just below the ribs." He then slowly throws the weight of his body forward to bear upon his own arms, and thus presses upon the *loins* of the subject and *indirectly on the abdominal contents*; the diaphragm is driven up into the chest, and air is forced out of the lungs. "This being effected, he gradually relaxes the pressure by bringing his own body up again to a more erect position, but without moving the hands; as he does this, air is drawn, by the removal of pressure and by their elastic reaction, into the lungs. The process is repeated quite regularly, and without manifest intervals between movements not less often than twelve times a minute; it may be done somewhat more rapidly, but fifteen times a minute would in any case be sufficient. By this means it is easily possible in an average man to effect an exchange of fully 6500 c.c. per minute—an amount which is more than enough to maintain complete aeration of the blood" (Schäfer).[2]

The prone position which is employed in this method prevents obstruction of the airway by mucus or by the falling back of the tongue.

The great merits of Schäfer's method are that it requires no apparatus whatever and that the technique is easily learnt.[3]

EVE'S ROCKING METHOD.—The patient is fixed on a stretcher; the head and feet are alternately tilted down to an angle of 45°; 8 or 9 rocking movements are carried out per minute. The 7 seconds available for each "respiration" should be divided into 4 seconds head down and 3 seconds feet down. In the head-down position the weight of the abdominal contents drives the diaphragm into the chest, producing expiration; in the feet-down position the abdominal contents fall away from the diaphragm, promoting inspiration. The method is easily carried out aboard ship (using a hammock) and is not fatiguing; it promotes the venous return and mechanically aids the circulation of the blood (p. 383).

The efficacy of these methods of artificial respiration has been determined on healthy volunteers who were deeply anæsthetized and then slightly overventilated to produce respiratory arrest; an endotracheal tube was

[1] Killick, Cowell, and Crowden, *Lancet*, 1939, ii, 897.

[2] Schäfer, *Proc. roy. Soc. Edin.*, 1905, *25*, 39.

[3] It is a matter of deep regret that so few medical men have received any training in how to perform the necessary movements with accuracy. Boy scouts, ambulance workers, and members of the police force are usually much better equipped to perform artificial respiration than doctors; Henderson has said that lives are not infrequently lost by physicians interfering with lay people who have been carrying out the emergency measures in an efficient manner; "it is easy to order a non-breathing victim of acute asphyxia into an ambulance, but he will be dead before he reaches hospital." It should be an important part of every physiological course to practise Schäfer's and the rocking methods of artificial respiration and by measuring the tidal air and pulmonary ventilation to demonstrate the efficacy of the procedures.

introduced to guarantee a free airway. The tidal air in two subjects was:

		A. 10 Stone.	B. 13 Stone.
		c.c.	c.c.
Schäfer's method:		340	530
Eve's method:			
On back:	30° rock	240	570
	45° rock	380	635
On face:	30° rock	340	725
	45° rock	580	850

The two methods have also been compared experimentally in dogs in which respiration was abolished by deep anæsthesia or section of the upper cervical spinal cord; in both groups muscle tone, including that of the diaphragm, was practically absent (as is the case in asphyxiated human subjects).[1] The volume of the tidal air was generally smaller with the Schäfer method, the mean advantage in favour of the rocking method being 50% and most pronounced when the loss of muscle tone was most complete. It may be concluded that though Schäfer's method is adequate, Eve's method may have advantages over it when a stretcher is available.

MECHANICAL METHODS.[2]—(1) *Drinker's Method.*—This involves the use of an airtight tank into which the patient is placed, with the head outside. Alternative negative and positive pressures are obtained in the tank by means of electrically driven pumps, and the effect is to produce movements of the chest wall resembling those of normal inspiration and expiration; the negative pressure pulls on the chest wall and produces inspiration; the positive tank pressure compresses the chest and produces expiration. Patients with respiratory paralysis following poliomyelitis have been kept alive for years in this way.

(2) *Bragg-Paul Method.*—A rubber bag is wrapped round the chest wall of the patient; by means of a pump positive pressure is applied at a given frequency to the chest. In this method, expiration is the active movement (as in Schäfer's method); when the pressure is released the chest passively enlarges and air is sucked in.

(3) In respiratory failure in the operating theatre artificial respiration can be given from an *anæsthetic machine*, from a simple pump like the *Oxford inflator*, or by blowing periodically down an inserted endotracheal tube.[3]

Whatever method is used it is essential to ensure that the airway is free.

Artificial respiration is called for in man in two types of respiratory failure:

(i) In the sudden cessation of breathing due to drowning, inhalation of irrespirable or poisonous gases, suicidal and accidental overdoses with narcotics, overdosage with anæsthetics or electrocution.

(ii) In gradually progressive respiratory failure due to paralysis of respiratory muscles, *e.g.* in anterior poliomyelitis or diphtheria.

In the *chronic* group there is usually ample warning of the impending disaster, during which time arrangements can be made to use a breathing

[1] Hemingway and Neil, *Brit. med. J.*, 1944, i, 833.
[2] Bourdillon *et al.*, *ibid.*, 1950, ii, 539.
[3] Electrical stimulation of the phrenic nerves in the neck has been carried out in man to produce rhythmic diaphragmatic contraction.

machine. In the *acute* group not a moment may be lost and the treatment must be capable of being carried out by the first instructed person on the scene and require the use of no more equipment than can be readily improvised almost anywhere. There is another fundamental difference between the chronic and acute forms of respiratory failure. In the former, machine-breathing is instituted while the *circulation is functioning normally*; if the mechanical respiration is efficient no circulatory derangement develops and the blood flow to all the organs remains normal. In acute asphyxia, on the other hand, failure of the circulation and of the central nervous system follows rapidly in the train of respiratory arrest. The body is unfortunately quite unadapted for dealing with complete or even very severe oxygen lack, which "stops the machine and wrecks the machinery" in a matter of minutes (p. 445). Let us consider the sequence of events in drowning, for example. After one minute or so of complete lack of oxygen consciousness is lost; within another minute or two the respiratory centre, which initially was stimulated, ceases to function. The vasomotor centre is more resistant and may maintain vasoconstriction for a little longer, but soon it too fails, and full peripheral vasodilatation sets in, with a resulting fall of blood pressure to about 40 mm. Hg. Most important of all, the heart—unlike skeletal or smooth muscle—can function normally for only a short time in the absence of oxygen; it has no type of anærobic metabolism to fall back on. The force of the heart beat in severe anoxia rapidly weakens and the chambers greatly dilate; the output into the blood vessels is reduced to a trickle, and the blood flow to the organs almost ceases. When fibrillation of the ventricles develops the chances of recovery are of the slenderest. Acute asphyxia thus presents a combination of respiratory, circulatory, and nervous system failure with which the treatment employed must cope effectively if it is to succeed.

There is complete agreement as to the immediate action which should be taken. The special conditions of the operating room were mentioned above; otherwise the Schäfer method should be instituted at once and no time lost while preparations are being made for the use of any other method. But it must be remembered that resuscitation demands more than just effective pulmonary ventilation; the oxygen introduced into the lungs must be transported by the circulation, especially to the brain, heart, and kidneys. As only a minimal circulation is present in asphyxiated patients, and life is being maintained most precariously, quite small differences in the circulatory state produced by the resuscitation technique may be of decisive importance. It is found experimentally that even with equal degrees of pulmonary ventilation the rocking method is associated with a greater cardiac output and therefore with a better blood flow to the organs, and with a higher level of oxygen consumption; these important advantages are largely due to obvious mechanical factors. In the head-down phase the venous blood drains out of the legs and abdomen into the heart, and the arterial blood flow into the head is facilitated. During the feet-down position venous blood drains out of the head, and the lower part of the body is fed with arterial blood, though the venous outflow from that region is impeded. Some of these advantages are lacking in the Schäfer method. A good case has been made out for changing over from the Schäfer to the rocking method as soon as suitable facilities are available.

REGULATION OF THE BREATHING [1]

Respiratory Centre.—The term *respiratory centre* is used here to denote the grey matter in the pons and upper medulla which is responsible for *automatic rhythmic* respiration. The centre so defined is subdivided into an *inspiratory* centre and an *expiratory* centre in the medulla and a *pneumotaxic* centre in the upper pons. The respiratory centre (predominantly its inspiratory centre) is *reflexly* regulated chiefly by afferent vagal impulses from stretch receptors in the lungs, by afferents from the chemoreceptors and the vascular pressure-receptors and from the higher levels of the brain. The inspiratory centre is extremely sensitive to changes in the chemical composition of the blood especially its CO_2 tension and H^+ ion concentration, and to a less extent its O_2 tension.

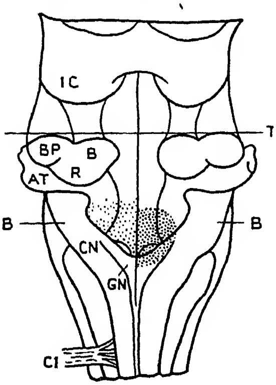


FIG. 227.—Localization of Inspiratory and Expiratory Centres. (Pitts *et al.*, *Amer. J. Physiol.*, 1939, *126*, 673.)

Dorsal view of brain stem (cat) with cerebellum removed. The respiratory centres are projected on the floor of the fourth ventricle.

To avoid overlapping the *expiratory* centre is shown only on the *left*, and the *inspiratory* centre only on the right.

Fig. 228 is a cross section at the level shown by the lines B–B. Section at level T, combined with double vagotomy, produces apneusis. The pneumotaxic centre thus lies in the pons above level T.

I.C., inferior colliculus.
B.P., brachium pontis (middle peduncle).
B., brachium conjunctivum (superior peduncle).
C1., first cervical root.
R., restiform body (inferior peduncle).
A.T., acoustic tubercle.
C.N., cuneate nucleus.
G.N., gracile nucleus.

INSPIRATORY CENTRE.—(i) The position of the inspiratory centre is mapped out by determining the region of the medulla which on stimulation produces inspiration. In the cat it extends from just below the upper border of the medulla down to the level of the calamus scriptorius (Fig. 227). It is scattered in the medial part of the *reticular formation*, especially in its more ventral part (Figs. 228, 229, A). The inspiratory centre and its descending path can be regarded as an afferent neurone leading to spinal motor neurones. As with stimulation of other afferent excitatory fibres, increased strength or frequency of stimulation applied to the inspiratory centre increases the depth of inspiration.

(ii) If the pneumotaxic centre is cut off by a trans-section through the lower pons, rhythmic breathing continues so long as the vagi are intact. If, however, the vagi are also divided, the inspiratory centre is said to be "*isolated.*" The inspiratory centre, when isolated in this way, is released from the dual inhibitory control of the pneumotaxic centre and the vagi (cf. pp. 386, 388). In the presence of an adequate CO_2 and O_2 tension the released inspiratory centre discharges steadily without intermission along descending paths to the spinal cord to stimulate the motor neurones supplying

[1] For general review, see Haldane and Priestley, *Respiration*, new edn., Oxford, 1935. Gray, *Pulmonary Ventilation and its Physiological Regulation*, Springfield, Ill., 1949, Pitts, Magoun, and Ranson, *Amer. J. Physiol.*, 1939, *126*, 673, 689; *127*, 654. Pitts, *J. Neurophysiol.*, 1942, *5*, 75, 403; 1943, *6*, 439; *Physiol. Rev.*, 1946, *26*, 609. Stella. *J. Physiol.*, 1938, *93*, 10, 263; *95*, 365.

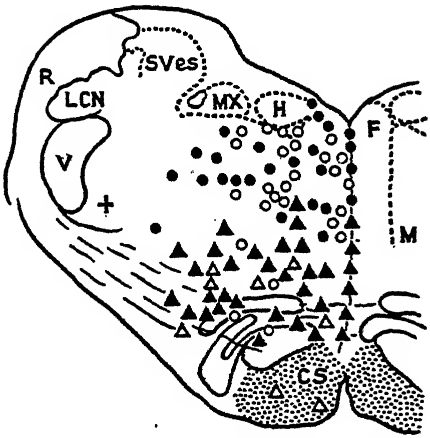


FIG. 228.—Distribution of Inspiratory and Expiratory Centres. (Pitts, *Amer. J. Physiol.*, 1941, *134*, 192.)

Section through medulla at level B–B, Fig. 227.
▲ Inspiratory centre.
○● Expiratory centre.
CS, Corticospinal (pyramidal) tract; F, Medial longitudinal fasciculus (bundle) H, Hypoglossal nucleus; LCN, Lateral cuneate nucleus; MX, Motor nucleus of vagus; M, Median lemniscus; R Restiform body SVes Vestibular nucleus; V, Spinal tract of V.

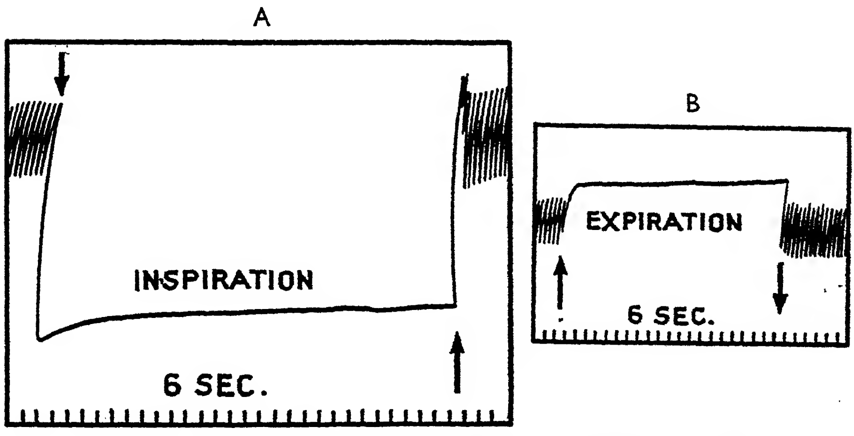


FIG. 229.—Effects of Stimulation of Inspiratory and Expiratory Centres. (Pitts *et al.*, *Amer. J. Physiol.*, 1939, *126*, 673.)

Downstroke = Inspiration.
Upstroke = Expiration.
A. Stimulation of Inspiratory Centre.
B. Stimulation of Expiratory Centre.

the muscles of inspiration. A prolonged powerful inspiratory effort (*apneusis*) thus develops (Fig. 230) which only comes to an end with the resulting anoxia becomes sufficiently severe to paralyse the activity of the cells of the centre.

EXPIRATORY CENTRE.—The expiratory centre can be similarly mapped out by determining the regions of the medulla which on stimulation produce expiration (Fig. 229, B). The cells of the expiratory centre are intimately intermingled with those of the inspiratory centre. In the cat they extend a little higher up in the medulla (Fig. 227) and lie more dorsally and more laterally in the reticular formation (Fig. 228). The descending fibres from the expiratory centre stimulate the spinal motor neurones which innervate the muscles of expiration. It is impossible to isolate the expiratory centre so as to produce a "release" of sustained expiratory activity; it is therefore presumed that the expiratory centre does not discharge spontaneously but only when it is appropriately stimulated.

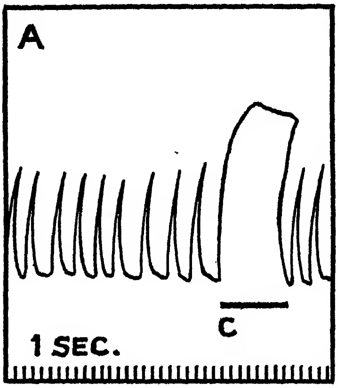


FIG. 230.—Spontaneous Discharge of Isolated Inspiratory Centre. (Stella, *J. Physiol.*, 1938, *93*, 268)

Cat. The brain stem has been divided in the upper pons to cut off the pneumotaxic centre. At the signal (c), both vagi were blocked by application of cold. As the inspiratory centre is released a powerful, sustained inspiration (apneusis) sets in. When the cold block of the vagi was removed, rhythmic breathing returned.

Reciprocal Innervation in Breathing.—The response to stimulation of the inspiratory and expiratory centres probably involves reciprocal innervation, *i.e.* when inspiratory spinal motor neurones are stimulated the antagonistic expiratory spinal motor neurones are reciprocally inhibited and vice versa (Fig. 235). Simultaneous stimulation of the inspiratory and expiratory centres leads to algebraic summation of their individual effects on the spinal motor neurones.

PNEUMOTAXIC CENTRE.—This centre is situated in the upper pons and periodically inhibits the inspiratory centre thus converting its spontaneous continuous discharge into a rhythmic pattern of discharge (inspiration) and rest (expiration).

(i) If the vagi are intact, cutting off the pneumotaxic centre still allows rhythmic breathing to continue.

(ii) If both vagi are cut, a trans-section above the upper border of the pons, which spares the pneumotaxic centre, likewise allows rhythmic breathing to continue.

(iii) If both vagi are cut and the pneumotaxic centre is *completely* cut off by a trans-section through the lower pons, the isolated respiratory centre sets up a deep inspiration which only ceases when respiratory failure, from anoxia, results. If the pneumotaxic centre is *partially* cut off, shorter bouts of apneusis, interspersed with brief gasps, appear and are repeated till the centre fails (Fig. 236).

It is suggested that normally when the inspiratory centre discharges downwards to the spinal cord, impulses also pass from it along collaterals to the pneumotaxic centre which is stimulated to send back *inhibitory* impulses to the inspiratory centre; when this inhibition reaches an adequate level the

inspiratory centre ceases to discharge, inspiration stops and expiration sets in. During expiration the pneumotaxic centre is no longer stimulated and so its inhibitory influence on the inspiratory centre is withdrawn; the latter centre resumes its spontaneous discharge producing the next inspiration; and so the cycle repeats itself. The inspiratory and pneumotaxic centres form a mutually reacting *closed circuit* which converts the steady discharge of the inspiratory centre into an alternating pattern of activity and rest (Fig. 235).

Rôle of the Vagi.—(1) Numerous vagal afferent filaments ramify in the walls of the alveoli of the lung; they are typical stretch receptors which are

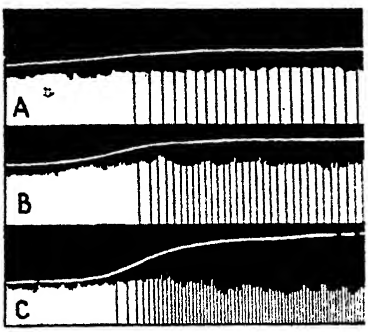


FIG. 231.—Vagal Action Potentials set up by Artificial Stretch of Lung.

Spinal Cat. Single fibre vagus preparation. The lungs are passively inflated by means of a pump. The movement of the white signal line is directly proportional to the degree of inflation.
A. Inflation=65 c.c. Maximum frequency of nerve impulses, 80 per sec.
B. Inflation=115 c.c. Maximum frequency 120 per sec.
C. Inflation=230 c.c. Maximum frequency 250 per sec. (Adrian, *J. Physiol.*, 1933.)

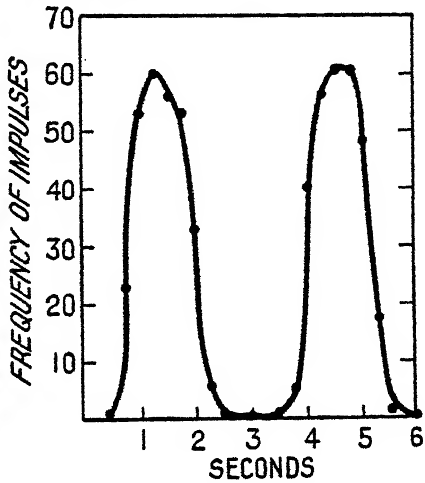


FIG. 232.—Vagal Action Potentials during Two Normal Respiratory Cycles.

Action potentials are recorded in a single fibre of the vagus nerve in the decerebrate cat. The frequency of the impulses reaches a maximum of 60 per sec. at the height of inspiration and falls to a minimum at the end of expiration. (Adrian, *J. Physiol.*, 1933.)

stimulated when the lungs are stretched during inspiration. Impulses (action potentials) can be recorded in the distal end of a *single* cut vagus fibre. The lungs are *passively inflated* quickly by means of a pump; impulses are set up which reach a peak frequency which is directly related to the degree of distension, *i.e.* the greater the degree of stretch of the lung, the higher the frequency of the impulses set up (Fig. 231). The pulmonary stretch receptors do not show adaptation, *i.e.* as long as the stretch is maintained the frequency of the discharge is kept up. With greater lung distension more vagal receptors are probably stimulated and thus more vagal afferents come into action. Increased stretch of the lungs thus results in more impulses arriving at the inspiratory centre in unit time.

(2) During *natural* breathing vagal impulses are set up with the onset of inspiration, and increase in frequency to a maximum as the lungs become

progressively filled with air ; with the onset of expiration the impulses begin to die down and disappear (Fig. 232).

(3) Evidence will now be presented to demonstrate that the vagal afferents (like the pneumotaxic centre but more powerfully) convert the steady discharge of the inspiratory centre into a rhythmic alternation of inspiration and expiration.

(i) Section of the vagi in the intact animal produces slower and deeper breathing (Fig. 236) ; the effect is smaller when one vagus only is cut (Fig. 233).

(ii) If the vagi are *intact*, a section through the upper pons to cut off the

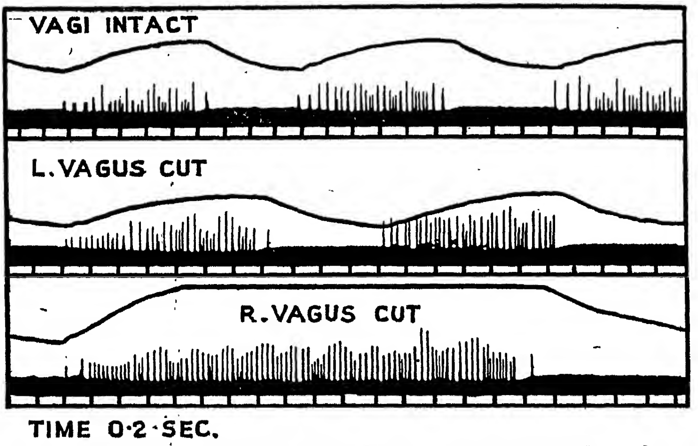


FIG. 233.—Effects of Section of Vagus Nerves in High Decerebrate Animal (Pneumotaxic Centre Intact). (Pitts, *J. Neurophysiol.*, 1942, *5*, 405.)

Impulses (action potentials) were recorded in about two fibres of the phrenic nerve representing (indirectly) the discharge of the inspiratory centre.
In each record, the upper tracing=respiration lower tracing=phrenic impulses.
Upstroke=inspiration.
Upper record=vagi intact.
Middle record=after section of the left vagus.
Lower record=after section of both vagi.
Note that after vagotomy, inspiration is more prolonged and its rate is reduced.

pneumotaxic centre leaves rhythmic breathing intact ; if one vagus is then cut, breathing becomes somewhat slower and deeper ; if both vagi are cut, the inspiratory centre is isolated, rhythmic breathing ceases and apneusis sets in.

(iii) The discharge of the *isolated* inspiratory centre can be conveniently studied by recording action potentials in a single fibre of the phrenic nerve ; the impulses follow one another steadily at a high frequency. If the central end of one vagus is stimulated there is after a latent period a decrease and finally an arrest of the inspiratory discharge ; on cessation of stimulation there is an inhibitory after-discharge, *i.e.* the inhibition of the inspiratory centre outlasts for some time the end of vagal stimulation, and then the discharge of the centre is resumed (Fig. 234).

The interpretation of these results is as follows : the *isolated* inspiratory centre is in a high state of central excitation and is vigorously generating

and discharging impulses. The afferent vagal impulses are inhibitory to this centre and produce in it the characteristic changes of reflex inhibition (p. 542). The latent period is due to the time taken for the impulses to reach the centre and for the central inhibition to increase sufficiently to overcome the existing state of central excitation; as more impulses arrive along the vagus, inhibitory recruitment occurs; more of the cells of the inspiratory centre come under the influence of the central inhibitory state until finally all activity is extinguished. On cessation of stimulation afferent vagal impulses continue to travel in complicated relay paths to reach the inspiratory

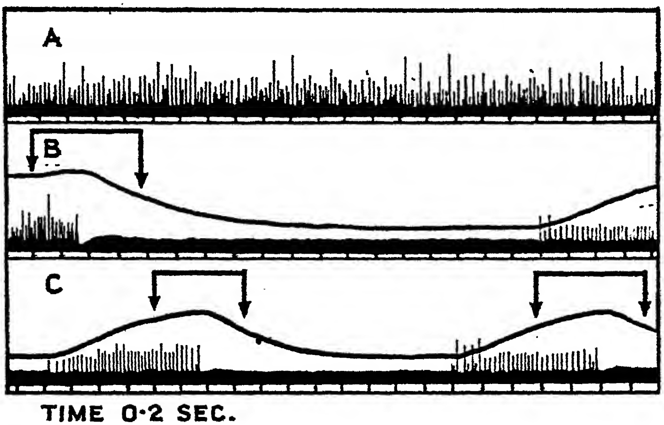


FIG. 234.—Inhibition of Isolated Inspiratory Centre by Stimulation of Central End of Vagus Nerve. (Pitts, *J. Neurophysiol.*, 1942, *5*, 407.)

The pneumotaxic centre has been cut off by a trans-section through the lower pons. Both vagi cut. Inspiratory centre isolated.

In each record, upper tracing shows respiration (upstroke = inspiration), lower tracing shows action potentials in a few phrenic fibres (representing discharge of inspiratory centre.)

Upper record (A): sustained discharge of isolated inspiratory centre (apneusis).

Middle record (B): Between arrows stimulate central end of vagus. After a latent period the inspiratory centre is arrested and expiration occurs. Note the long inhibitory after-discharge.

Lower record (C): Repeated stimulation of central vagus (betwe n arrows) cuts short inspiration and produces expiration.

centre and to maintain for some time the induced state of inhibition (inhibitory after-discharge). As the reverberation in the relay paths subsides, the inspiratory centre, freed from inhibition, resumes its discharge.

(iv) These results explain how the vagi function in the intact animal. During inspiration, the lungs are progressively stretched and more impulses pass up the vagi, until by inhibitory recruitment the discharge of the inspiratory centre is arrested and expiration sets in. The impulses passing up the vagi then die down and cease, but inhibitory after-discharge maintains the quiescence of the inspiratory centre producing a pause between the end of expiration and the next inspiration.

(v) The *deep rapid* breathing of muscular exercise or of CO_2 excess can be explained as follows. As the inspiratory centre is being stimulated directly and reflexly it is in a high state of excitation and discharges vigorously to the spinal motor neurones producing a powerful inspiration. The afferent

vagal impulses have to compete against this *higher* level of concentration; more afferent impulses than normally must arrive in unit time to produce this inhibition; inspiration thus proceeds longer, *i.e.* the inspiration is deeper, until with greater stretch of the lungs the afferent vagal backlash inhibits

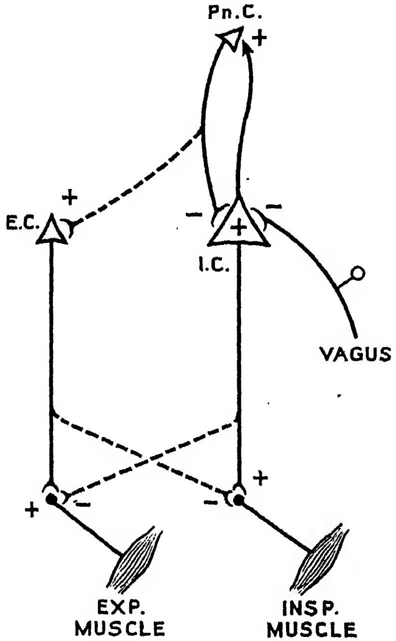


FIG. 235.—Diagram of Respiratory Centres.

Pn. C., Pneumotaxic Centre.
I. C., Inspiratory Centre.
The + sign in the cell indicates that the centre discharges spontaneously.
E. C., Expiratory Centre.
Insp. Muscle, Inspiratory muscle.
Exp. Muscle, Expiratory muscle.
Vagus, Vagal afferents from lungs.
The interrupted lines indicate the reciprocal innervation which probably takes place.
Thus: Pneumotaxic Centre inhibits Inspiratory Centre and may stimulate Expiratory Centre.
Inspiratory Centre stimulates spinal inspiratory motor neurones and inhibits expiratory spinal motor neurones.
Expiratory Centre stimulates expiratory spinal motor neurones and inhibits inspiratory spinal motor neurones.

the inspiratory centre. The inhibitory after-discharge likewise has to compete against a highly excited inspiratory centre; central excitation soon attains the upper hand and after a shorter interval the next inspiration starts. Breathing is thus both deeper and quicker. Expiration involves strong contraction of the expiratory muscles owing to powerful stimulation of the *expiratory* centre.

RHYTHMIC BREATHING (Fig. 235).—From what has been said above it is clear that the "inherent" steady discharge of the inspiratory centre is

converted into a rhythmic action by two automatically operating inhibitory mechanisms each of which is indirectly set into operation by the inspiratory centre itself:

(i) via the pneumotaxic centre (p. 386).

(ii) via a discharge to the inspiratory muscles, leading to inflation of the lungs, stimulation of the pulmonary stretch receptors and the setting up of afferent inhibitory impulses in the vagi (p. 387).

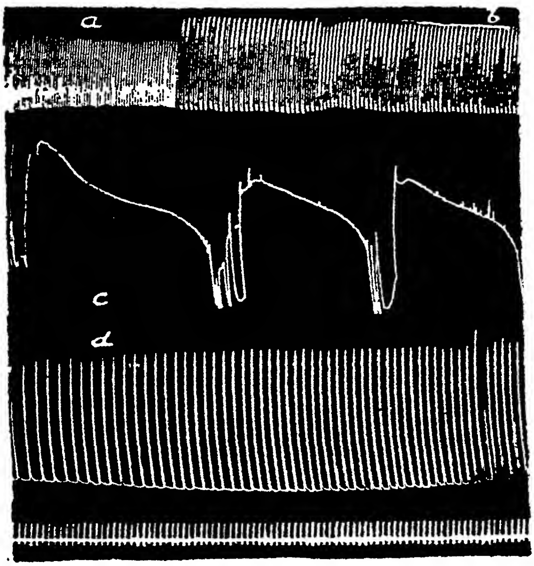


FIG. 236.—Effects of Injury to Central Respiratory Mechanisms.

Record of Respiratory Movements. Inspiration = Upstroke.

a, Normal; *b*, Slow, deep breathing following double vagotomy; *c*, Trans-section through pons in vagotomized animal, *partially* cutting off pneumotaxic centre. Apneustic breathing interspersed with gasps; *d*, Trans-section through upper medulla. Gasping breathing. (Lumsden, *J. Physiol.*, 1923, *57*).

EFFECTS OF INJURY TO RESPIRATORY CENTRES.—When the respiratory centre is depressed by severe asphyxia or anoxia, over-dosage with anæsthetics or medullary anæmia, normal breathing is replaced successively by apneuses and gasps and finally stops altogether (cf. Fig. 236, *c*, *d*). Conversely, if recovery is brought about by artificial respiration or other suitable methods, restoration of breathing occurs in the reverse order. Apneustic breathing may occur in infants following birth injuries; in such cases hæmorrhages have been demonstrated post-mortem in the pons.

Mechanism of Breathing.—The pattern of the nervous discharge from the normally controlled inspiratory centre can be conveniently studied by recording action potentials in single motor units of the respiratory muscles (p. 503) or in single fibres of the nerves supplying them (Fig. 237). The muscles

more muscle fibres are in action and a more extensive chest movement results. The expiratory muscles are treated in a similar manner.

When the inspiratory centre is depressed its discharge becomes of lower frequency, of shorter duration (Fig. 238, C) and involving fewer motor units; inspiration becomes progressively feebler and finally ceases.

Regulation of the Breathing.—Respiration is adjusted to serve a number of " purposes " in the body.

(1) *To supply oxygen and get rid of CO_2* formed in the body. Both the oxygen requirements and the amount of CO_2 liberated are proportional to

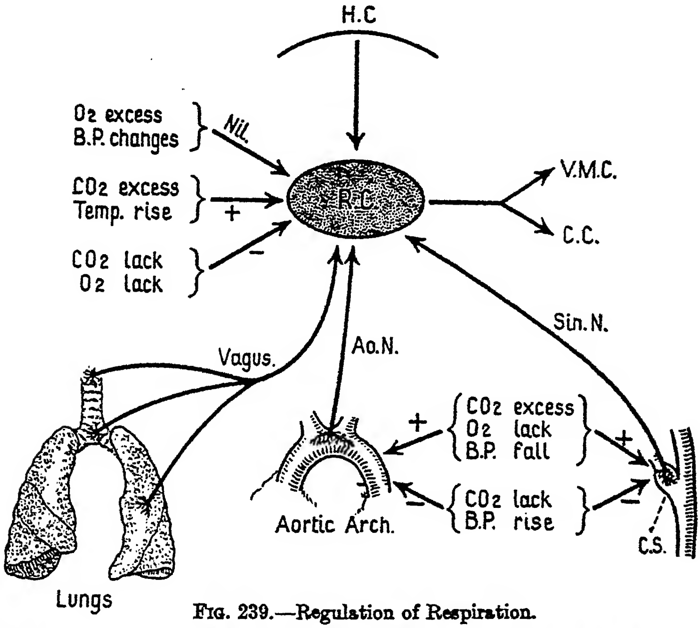


Fig. 239.—Regulation of Respiration.

R.C.=Respiratory centre; H.C.=Higher centres; C.C.=Cardiac centre; V.M.C.=Vasomotor centre; C.S.=Carotid sinus region (including carotid body); Sin.N.=Sinus nerve; Ao.N.=Aortic nerve (aortic arch includes the aortic body); +=Stimulates breathing; −=Depresses breathing.

the degree of activity of the body. All other things being equal, then, the *pulmonary ventilation is directly proportional to the metabolic rate.* The link between metabolism and breathing is probably the variation in the CO_2 tension in the blood. It should be emphasized that a level of ventilation sufficient to eliminate CO_2 in appropriate amounts (*at sea level*, and when *breathing air of normal composition*) is more than adequate to meet the oxygen requirements (except under conditions of extreme exertion).[1]

(2) To help to *regulate the H^+ ion concentration of the blood.* The breathing responds in an extremely sensitive manner to the slightest change in the H^+ ion concentration of the blood and in a way that tends to restore the

[1] " Over the oxygen supplies of the body CO_2 spreads its protecting wings " (Miescher).

reaction to its normal level. The *pulmonary ventilation thus varies with the reaction of the blood.*

(3) To maintain the appropriate *oxygen tension* of the blood in certain conditions of anoxia.

(4) To help to maintain the normal body *temperature.*

The main factors to be considered in our analysis are: (cf. Fig. 239):

(1) CO_2 excess and lack.

(2) Changes in H^+ ion concentration of blood.

(3) O_2 lack.

(4) Rise of blood temperature.

(5) Effects of changes in arterial blood pressure and cerebral blood flow.

(6) Voluntary and other reflex variations in breathing.

CO_2 and Breathing.—Fresh air contains negligible amounts of CO_2 (0·03%). When the inspired air contains excess CO_2 breathing increases first in depth and later in frequency, the former being quantitatively the more important. The pulmonary ventilation increases strikingly as shown in the Table below (from Haldane and Priestley). The venous blood as it

Percentage CO_2 in Inspired Air.	Average Depth of Respiration.	Average Frequency per Minute.	Ventilation of Alveoli with Inspired Air.	Percentage CO_2 in Alveolar Air.
0·04	673 c.c.	14	100 taken as normal.	5·6
0·79	739 c.c.	14	116	5·5
2·02	864 c.c.	15	153	5·6
3·07	1216 c.c.	15	226	5·5
5·14	1771 c.c.	19	498	6·2
6·02	2104 c.c.	27	857	6·6

flows through the lungs normally yields up a certain amount of CO_2 which passes into the alveoli whence it is eliminated by the fresh air breathed. If the inspired air contains much CO_2, it cannot eliminate CO_2 so well from the alveoli unless the pulmonary ventilation is changed. The alveolar CO_2 therefore rises slightly. As the arterial blood leaves the lungs in equilibrium with alveolar air, the arterial CO_2 tension also rises. Breathing is consequently stimulated, and so compensates for the inspired air being "contaminated" with CO_2. The alveolar CO_2 is thus reduced practically to its normal level. Alveolar CO_2 never falls quite to normal, for should it do so the stimulus to excessive breathing would cease and CO_2 would reaccumulate. In other words pulmonary ventilation is stimulated for as long as the CO_2-rich mixture is inhaled, thus keeping alveolar and arterial CO_2 tensions almost normal.

This compensatory mechanism has its limits, however. As the inspired CO_2 content approaches the level found in alveolar air it becomes increasingly difficult to eliminate CO_2 as rapidly as it is formed and (*e.g.* with inspired CO_2 of 5·14% the alveolar CO_2 rises (*e.g.* from 5·5 to 6·2%)). Obviously when the inspired CO_2% exceeds that in the alveolar air full compensation

is impossible and the alveolar CO_2 rises markedly (*e.g.* to 6·6% when the inspired CO_2 is 6·02%) in spite of tremendous overventilation. There is some mental confusion and depressed sensory acuity, the heart quickens and the blood pressure rises. (For effects on circulation cf. pp. 274, 309). When fresh air is breathed again, respiration soon tends to decrease towards normal, but some of the other symptoms may become temporarily aggravated (especially the headache) and in severe cases vomiting may occur. When the CO_2 concentration in the blood is very high, grave toxic effects appear, consisting of cardiac slowing, loss of consciousness, and then depression and finally cessation of breathing.

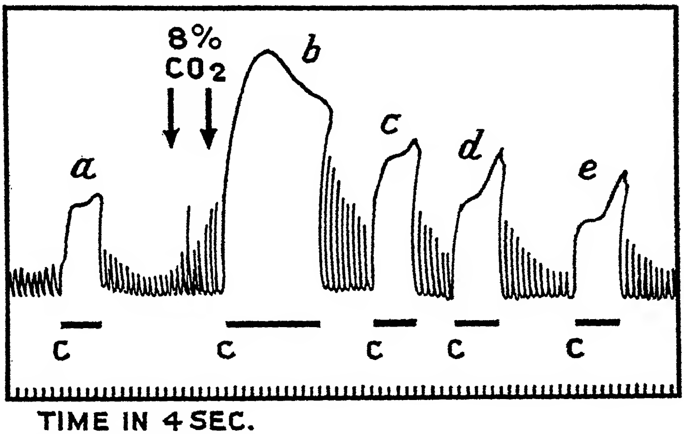


FIG. 240.—Effect of CO_2 Excess on Isolated Respiratory Centre. (Stella, *J. Physiol.*, 1938, *93*, 267.)

Cat. Trans-section through the upper pons to cut off the pneumotaxic centre. The inspiratory centre is temporarily isolated by blocking the vagi. The height of the resulting, sustained inspiration (apneusis) is a measure of the discharge of the inspiratory centre and, therefore, of its degree of excitation.

Control, breathing air. At C, block vagi. Note height of apneusis (*a*). Between arrows inhale 8% CO_2 in air. Note stimulation of breathing. At C, block vagi (including afferents from the aortic bodies). Note that the resulting apneusis (*b*) is greater than in (*a*). As the effect of CO_2 excess wears off, the height of the apneusis resulting from vagal block (C), *i.e.* (*c*), (*d*), (*e*), becomes progressively smaller.

MODE OF ACTION OF CO_2.—This is complex in character. CO_2 acts mainly *directly* on the inspiratory centre and partly *reflexly* by stimulating chemoreceptors in the carotid and aortic bodies (p. 738), thus sending excitatory impulses to the centre.

(i) *Direct Action.*—In an animal with a low pontine trans-section (to cut off the pneumotaxic centre), the two vagi are blocked by means of cold; the released, isolated, *inspiratory* centre sets up an apneusis and its height is noted (Fig. 240 (*a*)). The vagi are warmed and rhythmic breathing returns. The animal is given a CO_2-rich mixture to breathe and during the response the vagi are blocked again; the resulting apneusis is far bigger than in the control experiment (Fig. 240 (*b*)); this observation proves that CO_2 stimulates the inspiratory centre. After denervation of the carotid and aortic bodies, CO_2 still stimulates the isolated inspiratory centre (though to a smaller extent),

proving that its action is in part a *direct* one. Conversely, if the vagal block is carried out after a bout of overventilation to eliminate CO_2 the resulting apneusis develops very gradually and is feeble (Fig. 241). As results, qualitatively similar through quantitatively smaller, are obtained after denervating the carotid and aortic bodies, CO_2 lack depresses the inspiratory centre in part, by a *direct* action.

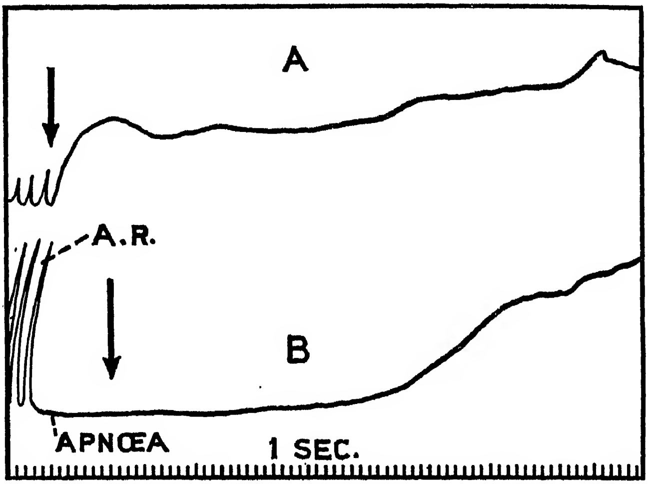


FIG. 241.—Effect of CO_2 lack on Isolated Inspiratory Centre. (Stella, *J. Physiol.*, 1939, *93*, 269.)

Cat. Trans-section through the upper pons to cut off the pneumotaxic centre. Isolate the inspiratory centre temporarily by blocking the vagi: the height of the resulting, sustained inspiration (apneusis) is a measure of the discharge of the inspiratory centre and, therefore, of its degree of excitation.

A. Control: quiet breathing. At arrow, block vagi. Note rate of developmen and height of apneusis.

B. Artificial respiration (A.R.) has been carried out to induce CO_2 lack. When it was stopped, apnœa set in.

During the apnœa, at the arrow, block the vagi (including afferents from the aortic bodies). Note the long latent period before the apneusis begins to develop.

(ii) *Reflex Action.*—The reflex effects of CO_2 via the chemoreceptors are discussed on p. 745.

H+ ion Concentration and Breathing.—Breathing is very sensitive to changes in H+ ion concentration of the blood; the pulmonary ventilation changes in a manner that tends to restore the blood reaction to its normal value. As was pointed out above (p. 393), one of the main functions of breathing is to help to maintain the blood reaction within proper limits.

RESPONSE TO ACIDÆMIA.—When the H+ ion concentration of the blood tends to rise, the pulmonary ventilation is increased, more fresh air is taken into the lungs, more CO_2 is eliminated, and the level of the alveolar CO_2 falls below normal. As the alveolar CO_2 tension regulates the amount of H_2CO_3 in the arterial blood, a fall of alveolar CO_2 results in a lowered arterial CO_2

and H_2CO_3 content. The numerator in the expression $H_2CO_3/BHCO_3$ which determines the H^+ ion concentration (cH) of the plasma is decreased and thus the blood reaction tends to return to normal (p. 92).[1] The tendency to acidæmia which initiated these changes is thus combated (cf. Fig. 57).

Such a response occurs in the following conditions:

(i) During the secretion of the alkaline digestive juices (pancreatic juice, bile, and succus entericus) as sodium bicarbonate is excreted leaving a relative excess of H^+ ions in the blood.

(ii) A meat diet gives rise to excess of acid radicals, especially H_3PO_4 and H_2SO_4 derived from the oxidation of the P and S of the protein molecule, and therefore tends to cause an acidæmia (cf. p. 99).

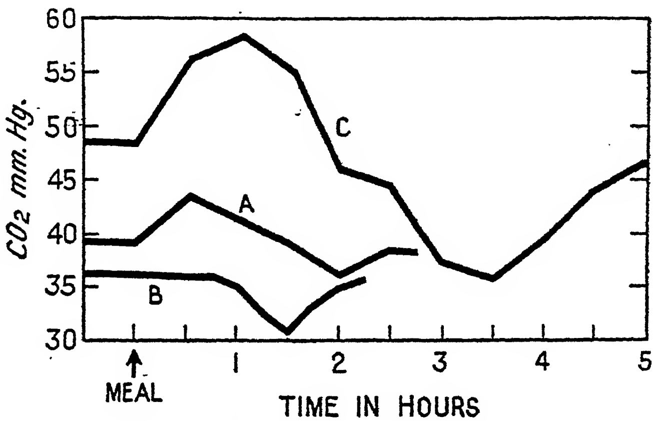


FIG. 242.—Changes in Alveolar CO_2 Tension with Meals.
Ordinate = Alveolar CO_2 tension in mm. Hg. A = Normal subject; B = Healthy subject with complete achylia gastrica; C = Case with marked hypersecretion of gastric juice. (Dodds and Bennett, *J. Physiol.*, 1921.)

(iii) If ammonium chloride is ingested, the NH_3 portion is split off; it is combined with carbonic acid and converted into urea.

$$NH_4Cl \rightarrow \underset{\substack{\downarrow \\ \text{urea}}}{NH_3} + H^+ + Cl'$$

H^+ ions are liberated this tends to produce acidæmia (cf. p. 98).

(iv) In diabetes mellitus and starvation, β-hydroxybutyric and aceto-acetic acid accumulate in the blood (p. 924).

In all these cases of tendency to acidæmia, increased breathing occurs with resulting lowering of the alveolar and arterial CO_2 tension.

RESPONSE TO ALKALÆMIA.—When the H^+ ion concentration of the blood tends to fall, the pulmonary ventilation is diminished, CO_2 is retained, the alveolar CO_2 consequently rises, the arterial CO_2 and H_2CO_3 rise, the H^+ ions in the blood are increased and the normal blood reaction is restored. The increase in H_2CO_3 in the expression $H_2CO_3/BHCO_3$ compensates for the

$$pH = (-\log_{10} cH) = 6{\cdot}1 + \log \frac{BHCO_3}{H_2CO_3}.$$

initial increase in $BHCO_3$ (cf. Fig. 57). These general principles are well illustrated by the following examples:

(i) During gastric secretion HCl is eliminated from the blood, with a consequent tendency to alkalæmia.

(ii) The same effect is produced by the ingestion by mouth of $NaHCO_3$, which is absorbed as such into the blood stream.

(iii) A vegetable diet contains a relative excess of basic radicals, and so tends to increase the alkalinity of the blood.

In all three conditions the pulmonary ventilation is diminished and the alveolar CO_2 rises. As a result, the arterial CO_2 tension, the H_2CO_3 content and the concentration of H^+ ions in the blood all rise, thus compensating for the initial alkalæmic tendency.

Fig. 242 shows the compensatory changes in the alveolar air (and therefore in the arterial blood) occurring with meals. The initial rise of alveolar CO_2 runs parallel with the amount of gastric acid secreted. In achlorhydria—in which condition no gastric HCl is formed—no initial change is observed in the alveolar air. During the phase of alkaline secretion (*i.e.* of pancreatic juice and bile) the alveolar CO_2 falls. These changes are best seen after the first meal of the day; after later meals the effects are not well marked, because the response to one meal blurs the reactions to the succeeding one.

Mode of Action of H^+ ion Changes.—As in the case of CO_2, changes in the H^+ ion concentration affect breathing in two ways: (i) by stimulating the respiratory centre directly; (ii) reflexly, by acting on the chemoreceptors in the carotid and aortic bodies (p. 745).

Effects of Oxygen Lack on Breathing.—The effects of oxygen *lack* are complex and depend in part on the degree of acuteness and severity of the deprivation.

(1) If an *inert gas like nitrogen* is breathed, the CO_2 is eliminated quite normally from the alveoli, but the oxygen present there is washed out as well. Not only does the venous blood reaching the lungs soon cease to gain any oxygen, but when the oxygen content of the alveoli falls sufficiently, the *venous blood gives up its oxygen* and the arterial blood leaves completely reduced. Some initial hyperpnœa occurs as a rule, but within 45 seconds loss of consciousness develops suddenly, and practically without any warning, from oxygen lack to the brain. There are no reserve stores of oxygen in the tissues (unlike CO_2); the only oxygen present there is the minute amount which is present in solution. Soon after, breathing decreases and ceases from failure of the respiratory centre.

(2) If the *oxygen lack is less severe*, the effects produced depend on how rapidly it is induced.

If the subject breathes in and out of a large bag through a soda-lime tower (to absorb CO_2) the CO_2 evolved is removed and does not therefore accumulate in the bag, but the oxygen content gradually diminishes. The breathing is unaffected till the oxygen in the bag is reduced by about one-third, *i.e.* from 21% to 14%. In sensitive subjects a considerable increase in the breathing may now occur; but often the increase is slight. It contrasts very strikingly with the early intense hyperpnœa which results from even a small rise of CO_2 in the inspired air. The effects on respiration and circulation in a typical experiment are set out graphically in Fig. 243. The changes in the alveolar air are shown in the Table on p. 400.

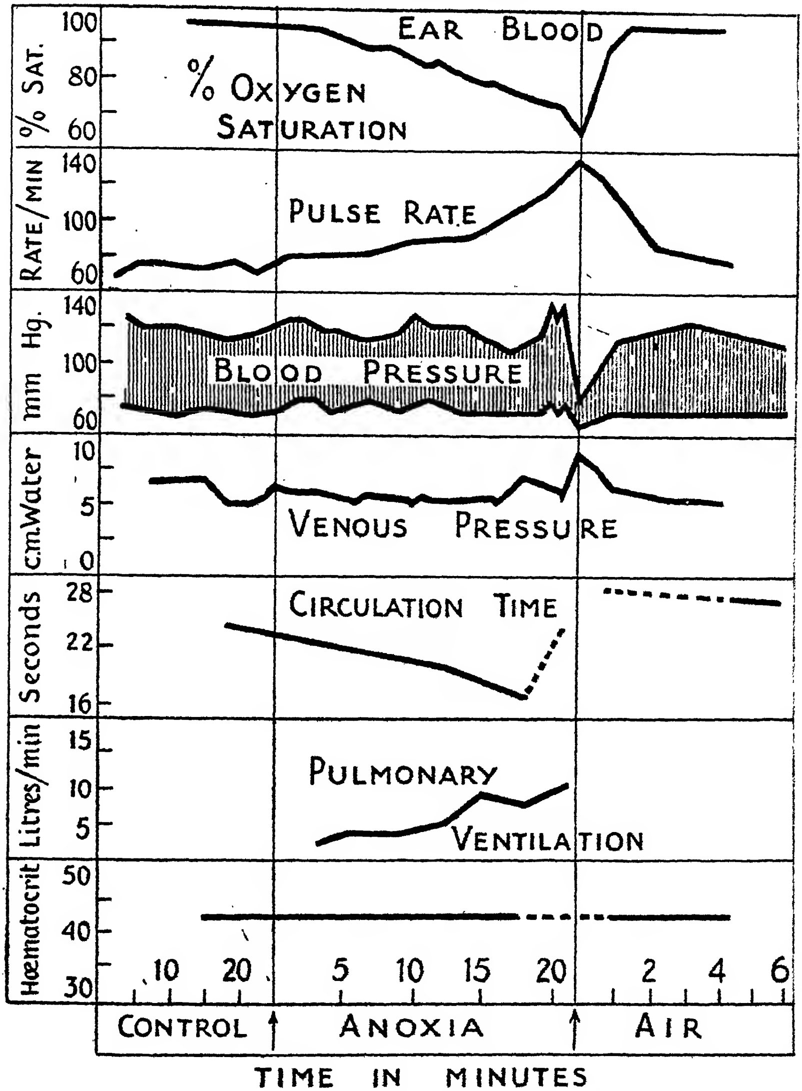


FIG. 243.—Circulatory and Respiratory Responses to Progressive Anoxia in Man.

After a control period the subject rebreathed from a bag containing air, through soda lime to absorb CO_2; the experiment was terminated when the subject fainted. The percentage saturation of the arterial blood with oxygen fell from 95 to 65; the pulmonary ventilation gradually increased from 4 to 10 litres per minute; the circulation time initially fell from 24 to 17 seconds (suggesting increased cardiac output); the pulse rate rose from 65 to 130 per minute. In the blood pressure record, upper line = systolic, lower line = diastolic, shaded area = pulse pressure. The blood pressures remained unchanged for the first 20 minutes of anoxia. Just before consciousness was lost, systolic and diastolic blood pressure fell abruptly, the venous pressure rose, and the circulation time was lengthened, indicating acute heart failure. On breathing air normal conditions were restored. The hæmatocrit value was unchanged throughout; *i.e.*, there was no change in red cell concentration and no evidence of discharge of stored red corpuscles. (Redrawn from Erahler *et al.*, *Amer. J. Physiol.*, 1943, *138*, 595.)

Inspired Air O_2%	Alveolar Air O_2%	Alveolar Air CO_2%	Pulmonary Ventilation.
20·9	14·5	5·54	Control
16·0	10·4	5·62	Unchanged
12·8	8·3	5·37	Slightly increased
11·1	7·1	4·89	Increased
6·2	4·3	3·57	Greatly increased
63·7	57·6	5·41	Unchanged
80·2	72·2	5·84	Unchanged

Normally the alveolar O_2 percentage (14%, 100 mm. Hg pressure) is about 7% less than the atmospheric O_2 percentage (21%, 150 mm. Hg. pressure). The effect of the increased breathing is to make better use of such oxygen as is available in the inspired air; the difference between the O_2 concentration in the atmosphere and the alveolar air is less than normal. The blood is therefore better oxygenated than it would have been in the absence of the compensatory hyperpnœa. At the same time CO_2 elimination is increased; as there is no corresponding increase in CO_2 production the alveolar CO_2 tension falls. The associated fall in arterial CO_2 tension depresses the central respiratory mechanism and so prevents the anoxia from exerting its full stimulating action on breathing. For these reasons also there is a tendency, too, for breathing to wax and wane, *i.e.* be *periodic* in character. The effects may be summarized thus:

Oxygen lack—stimulates breathing—increases pulmonary ventilation—lowers alveolar CO_2—lowers arterial CO_2—alkalæmia—depresses breathing.

It is thus difficult to establish adequate compensation, as the stimulus of oxygen lack to breathing constantly tends to be counteracted. Cyanosis develops when the oxygen content of the inspired air falls to 10%; there is then marked mental incapacity. Consciousness is lost a little later, from oxygen deprivation of the higher centres.

(3) The results are different when *oxygen lack is of very gradual onset* and ample time is available for all the compensatory activities of the body to be brought into play. This is well seen on ascending *gradually* to high altitudes. Fig. 244 shows the relationship between altitude and barometric pressure. The percentage composition of the atmosphere is unaffected by altitude, but the *oxygen pressure* in the inspired air is reduced proportionately to the decrease in total atmospheric pressure.

Symptoms set in when a height of about 11,500 feet is reached which corresponds to a barometric pressure of about 500 mm. Hg, *i.e.* two-thirds of the normal level. With rapidly developing anoxia the chief stumbling-block in the way of an effective respiratory response is the excessive washing-out of CO_2 from the blood, which tends to depress breathing. It is found, however, that after staying for some time in rarified atmospheres a process of *acclimatization* develops and persistent hyperpnœa is maintained, in spite of the lowered alveolar (and arterial) CO_2 tension. Thus in acclimatized subjects on Pike's Peak in Colorado the pulmonary ventilation had increased from 10·4 to 14·9 litres per minute, and the alveolar CO_2 tension was 27 mm.

Hg instead of the normal 40 mm. Hg. Warren, during the 1938 Everest[1] expedition, found that at 22,000 feet (barometer = 340 mm.) the alveolar CO_2 tension was 17 mm.; (the alveolar oxygen tension was 40 mm. !). The breathing at these great heights was very rapid, about 55 to the minute, and very deep. At 28,000 feet, Somervell found that for every single step forward and upward 7–10 complete respirations were required. Values for the pulmonary ventilation are not given, but it must have been very great. Breathing quickly and deeply was very easy at high altitudes owing to the low density of the air.

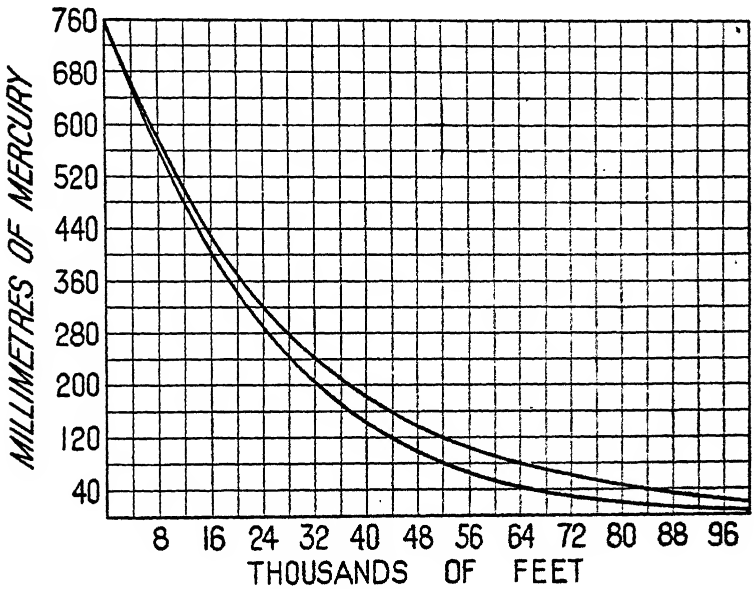


FIG. 244.—Curves showing Relationship of Barometric Pressure to Altitude. (Haldane and Priestley, *Respiration*, 1935.)

The upper curve is calculated from the formula of Zuntz and others, assuming a mean temperature of 15° C. The lower curve is calculated according to the internationally adopted system of altimeter graduation.

To enable such an enormous ventilation to continue in spite of the low arterial CO_2 tension, it is obvious that the alkalæmia must be compensated for in some way. This is done by the kidney, which excretes a less acid urine and less ammonium salt; the urine may even become alkaline. The kidney thus eliminates the excess of base, keeps the blood reaction fairly constant, and permits oxygen lack to continue to stimulate ventilation (cf. p. 98).[2]

The sequence of events may be thus summarized: O_2 lack—increases pulmonary ventilation—alkalæmia—compensated for by the kidney.

As a rule in acute experiments the increase in breathing due to oxygen

[1] But according to Mr. Chiu, in 1717 the mountain was first given the Chinese name of Chu-mu-lang-ma (Peak of Sacred Mother's Spring).

[2] For other effects of oxygen lack, see pp. 442 *et seq.*

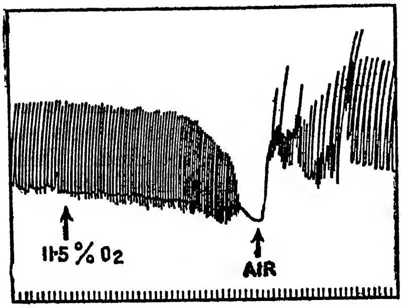


FIG. 245.—Effects of Oxygen Lack on Respiration. (Selladurai and Wright, *Quart. J. exp. Physiol.*, 1932.)

Cat. Decerebrate. Sinuses denervated and vagi cut. Time in 5 sec. Inspiration = upstroke. Inhalation of 11·5% O_2 mixture at arrow. Rapid onset of respiratory failure. Air given at second arrow. Breathing gradually recovers.

lack is chiefly in rate and not in depth, thus contrasting with the effects of CO_2 excess.

MODE OF ACTION OF OXYGEN LACK.—Moderate oxygen lack stimulates breathing in the intact animal. But if the *sinus nerves and vagi are severed*, inhalation of the same oxygen-poor mixture usually produces a rapidly developing *paralysis of respiration* (Fig. 245). Oxygen lack thus *directly depresses the "deafferented" respiratory centre*; suitably performed perfusion experiments show that anoxia acts on the sino-aortic chemoreceptors, setting up excitatory impulses which *reflexly stimulate* breathing (cf. p. 745). As a rule, in the intact animal the reflex stimulation prevails and increased ventilation results. It was explained on p. 400 how this hyperpnœa by washing out CO_2 sets up further complications.

The action of anoxia on the *inspiratory* centre can be demonstrated in the same way as described for CO_2 (p. 395). The pons is trans-sected to cut off the pneumotaxic centre; when the vagi are temporarily cold-blocked, the nspiratory centre is "isolated"; it discharges steadily and a control apneusis

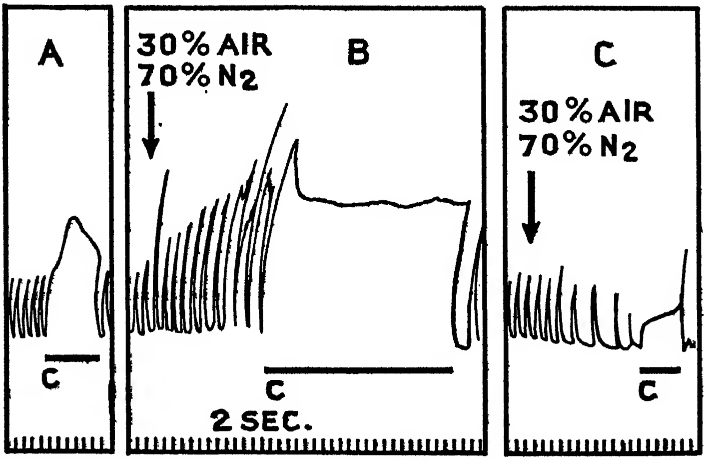


FIG. 246.—Effect of Oxygen Lack on Inspiratory Centre. (Stella, *J. Physiol.*, 1939, *95*, 366.)

Cat. Trans-section through the upper pons to cut off the pneumotaxic centre. On blocking the vagi (*c*), the inspiratory centre is released. The height of the resulting inspiration (apneusis) is a measure of the activity of the inspiratory centre.

A. Control, breathing air; block vagi: note height of apneusis.

B. At arrow, inhale 6% O_2. Breathing is reflexly stimulated; block vagi (*c*): note more powerful apneusis than in A.

C. Denervate carotid sinuses. Inhale 6% O_2. Breathing is depressed. Block vagi (including afferents from aortic bodies) (*c*): note very feeble apneusis. Oxygen lack directly depresses the inspiratory centre.

is recorded (Fig. 246, A). The vagi are warmed again, thus allowing rhythmic breathing to return. The animal is given an O_2-poor mixture to breathe and at the height of the response the vagi are again cold-blocked; a more powerful apneusis sets in, showing that anoxia stimulates the inspiratory centre (Fig. 246, B). The stimulation might be a direct or a reflex one. The experiment is repeated after cutting the sino-aortic nerves coming from the chemoreceptors; when the vagi are blocked again during a period of anoxia a smaller apneusis than in the control develops (Fig. 246, C). Anoxia thus *directly depresses* the inspiratory centre, but, via the chemoreceptors reflexly stimulates it.

Effects of Oxygen Excess.—Pure oxygen (at *atmospheric pressure*) can be breathed in man without ill effect for short periods, *e.g.* for several hours. There are few records of the effects of such inhalations for longer periods owing to fear of producing the pulmonary damage which develops in small animals breathing pure oxygen for several days; 100% O_2 has been breathed in a few instances for two days without ill effect in man; 60% O_2 *mixtures can be breathed in man indefinitely with perfect safety.* It is necessary to emphasize these facts owing to the very widespread confusion on the subject. There is *no* change in pulmonary ventilation, metabolic rate or blood pressure; mental activity is *not* stimulated. As pointed out on p. 194, erythropoiesis in the red marrow is depressed. The cerebral blood flow is slightly decreased owing to local vasoconstriction (p. 306).

If men are exposed to 4 *atmospheres* oxygen pressure, grave symptoms develop in less than 1 hour; these consist of faintness, fall of blood pressure, or convulsions.[1] In a series of healthy men exposed to 3·7 atmospheres O_2 pressure, great and unpredictable variations in tolerance were observed, symptoms occurring in 6–96 minutes. A very common early symptom was lip-twitching; twitching also occurred in the arms; other symptoms were dazzle, giddiness, nausea, vomiting, irregular respiration, confusion and in some instances convulsions and unconsciousness.[2] It has been proved by experiments on isolated brain slices that high O_2 pressure directly poisons many of the enzymes concerned with tissue oxidation thus paradoxically inducing cerebral anoxia.[3] The symptoms of oxygen poisoning are mainly referable to the central nervous system which is highly susceptible to the effects of anoxia.[4] High oxygen pressure may also decrease the cardiac output and cause cerebral vasoconstriction.

Climatic Conditions and Breathing.—When the body is exposed to warm and very moist external surroundings, regulation of body temperature becomes very difficult, as heat loss by evaporation and radiation from the skin is greatly interfered with. The pulmonary ventilation under these

[1] Behnke *et al.*, *Amer. J. Physiol.*, 1935, *110*, 565. Bean, *Physiol. Rev.*, 1945, *25*, *1*.

[2] Donald, *Brit. med. J.*, 1947, i, 667, 711.

[3] Dickens, *Biochem. J.*, 1946, *40*, 145, 171.

[4] The poisonous effects of high O_2 pressures have been attributed in part to *retention of CO_2 in the body.* At 2 atmospheres O_2 pressure the amount of O_2 dissolved in the plasma is over 4 c.c.% (normal 0·3 c.c.); this is sufficient to meet the resting O_2 requirements of the tissues (4 c.c. per 100 c.c. of blood flow); consequently no reduction of oxyhæmoglobin takes place. As this substance is more acid than reduced hæmoglobin, base is not made available for the uptake of the CO_2 formed in the tissues (cf. Fig. 253); oxyhæmoglobin also combines less readily than does reduced hæmoglobin with CO_2 to form carbamino-hæmoglobin (Fig. 252).

conditions is considerably increased; more water vapour is eliminated in the expired air, thus getting rid of extra heat from the body. The alveolar and arterial CO_2 tension falls; the resulting alkalæmia is compensated for in the usual way by the excretion of an alkaline urine (cf. p. 478, and Fig. 56, p. 99).

Blood Pressure, Cerebral Circulation, and Breathing.—A *sudden marked* rise of blood pressure in the resting animal depresses respiration; a fall of blood pressure stimulates breathing. These effects are reflexly produced by the presso-receptors in the carotid sinus and aortic arch. The evidence is as follows:

(i) A rise of pressure in the perfused carotid sinus reflexly inhibits breathing (Fig. 477, p. 741); conversely, a fall of carotid sinus pressure (*e.g.* as a result of occluding both common carotid arteries) reflexly stimulates breathing (Fig. 480, p. 743). By appropriate methods, similar results can be obtained by modifying aortic blood pressure; the afferents concerned are the endings in the adventitia of the aortic arch.

(ii) The increased breathing characteristic of hæmorrhage (p. 83) is produced reflexly partly by the fall of blood pressure (and partly by the anoxia).

(iii) The rise of blood pressure set up by adrenaline is accompanied in the intact animal by depression or cessation of breathing (Fig. 458, p. 726). This adrenaline apnœa is almost wholly reflexly produced by the rise of blood pressure acting on the aortic and sinus nerve endings. The apnœa cannot usually be brought about after section of the vagi and sinus nerves, *i.e.* a rise of blood pressure does not *directly* depress the respiratory centre.

It should be remembered, however, that in *intact man* a rise of blood pressure is frequently (in fact, commonly) associated with *increased* pulmonary ventilation. Thus, in vigorous exercise the breathing is stimulated in spite of the rise of blood pressure because the chemical changes in the blood (*e.g.* increased CO_2 and H^+ ion concentration) overcome the inhibitory reflex from the presso-receptors.

The respiratory centres, more than the other bulbar centres, rapidly lose their excitability if their oxygen supply is cut off (Fig. 245) or if their blood supply is interfered with, as by raised intracranial pressure (p. 464) or by cerebral anæmia.

Rôle of Afferent Nerves in Breathing.—(1) VAGI.—(i) The rôle of the vagi in normal respiration is fully discussed on pp. 387 *et seq.*

(ii) The afferent fibres from the laryngeal mucous membrane are concerned with the *cough reflex* which guards the respiratory passages against the entrance of foreign bodies.

(2) AFFERENTS FROM THE CARDIO-VASCULAR SYSTEM AND CHEMORECEPTORS.[1]—(i) It was shown above that blood pressure changes reflexly modify breathing by acting on the nerve endings in the aortic arch and carotid sinus.

(ii) A rise of venous pressure reflexly stimulates breathing (cf. pp. 273, 461).

(iii) The reflex effects of variations in CO_2 tension and H^+ ion concentration, of O_2 lack, and of various drugs, on respiration are considered on p. 745.

[1] Heymans *et al.*, *Arch. int. Pharmacodyn*, 1927, *33*, 273; 1930, *29*, 440; *J. Physiol.*, 1930, *69*, 254; *Le Sinus Carotidien*, 2nd edn., Paris, 1933. Schmidt and Comroe, *Physiol. Rev.*, 1940, *20*, 115.

These results must be considered in conjunction with those described in the discussions of the mode of action of CO_2, H^+ ions, and oxygen lack on breathing (pp. 394, 396, 398). The entire evidence proves that the chemical changes in the blood regulate breathing to a *significant extent reflexly from the chemoreceptors in the carotid and aortic bodies.*

(3) OTHER FACTORS INFLUENCING BREATHING.—Breathing is also influenced by the following factors:

(i) the *higher centres, e.g.* cerebral cortex or hypothalamus (Figs. 433, 449);

(ii) *afferent impulses* from the surface of the body, *e.g.* painful or thermal stimuli;

(iii) during *swallowing*, breathing is reflexly inhibited by impulses coming in the glossopharyngeal nerve from the post-pharyngeal wall;

(iv) during *sleep* the respiratory centre is depressed; although the metabolic rate is reduced, the respiratory centre cannot maintain a level of pulmonary ventilation which is adequate to deal even with the reduced rate of CO_2 formation; the alveolar CO_2 tension consequently rises considerably.

General Survey of Control of Breathing.—(1) In the resting individual the pulmonary ventilation is *adjusted to the rate of metabolism.*

The basal metabolism in an average man is about 70 Cal. per hour (p. 378), or 1·15 Cal. per minute. The O_2 intake per Cal. is 210 c.c. (as 1 litre of O_2 burnt in the body yields about 4·8 Cal. (p. 375)). The O_2 requirement under basal conditions is thus about 240 c.c. per minute.

From every 100 c.c. of inspired air about 4 c.c. of oxygen are absorbed into the blood. So the *pulmonary ventilation* at complete rest is $\frac{240}{4} \times 100 =$ 6 litres per minute. In a woman the basal pulmonary ventilation is about 4·5 litres per minute.

It is clear from the above considerations that any of the factors (enumerated on p. 378) which modify the metabolic rate produce corresponding alterations in the pulmonary ventilation. The effects of different *degrees of exertion* on the respiratory exchanges are well shown in the Table below (from Haldane and Priestley). The effects of exercise are fully discussed on pp. 435 *et seq.*

	O_2 intake per min.	CO_2 output per min.	Pulmonary Ventilation per min.	Alveolar CO_2%	R.Q.
Rest in bed	237 c.c.	197 c.c.	7·7 litres	5·97	0·88
Rest, standing	328 c.c.	264 c.c.	10·4 litres	5·70	0·80
Walking—					
2 miles per hour	780 c.c.	662 c.c.	18·6 litres	6·04	0·85
3 miles per hour	1065 c.c.	922 c.c.	24·8 litres	6·10	0·86
4 miles per hour	1595 c.c.	1398 c.c.	29·0 litres	6·36	0·88
5 miles per hour	2543 c.c.	2386 c.c.	60·9 litres	6·10	0·90

The link between metabolism and breathing is mainly the *alteration in the CO_2 tension of the blood,* to which breathing is exquisitely sensitive.

(2) Breathing is comparatively insensitive to *oxygen lack,* unless it is severe. Therefore at sea level, when air of normal composition is inspired,

oxygen lack plays little part in controlling breathing. At high altitudes, or in oxygen-poor atmospheres, it becomes the dominating factor and stimulates breathing, *although there is no change in metabolic rate,* and *although the effect of the hyperpnœa is to produce an alkalæmia.* The "goal" of the hyperpnœa is, of course, to *improve the oxygenation of the blood.*

(3) Changes in *blood reaction* due to normal bodily processes (as discussed on p. 397) are constantly modifying the pulmonary ventilation above or below that required by the metabolic rate in order to preserve the optimum H^+ ion concentration.

(4) External moist *heat* sets up hyperpnœa, which tends to cool the body; incidentally an alkalæmia is produced which has to be compensated for in appropriate ways (p. 478).

(5) The *level of the blood pressure* probably only affects breathing reflexly in the intact individual when it is altered rapidly and to a significant extent (p. 404) and if the subject is at rest.

Effects of Voluntary Hyperpnœa.—The subject breathes very deeply and quickly for 2–3 minutes and then allows his breathing to act independ-

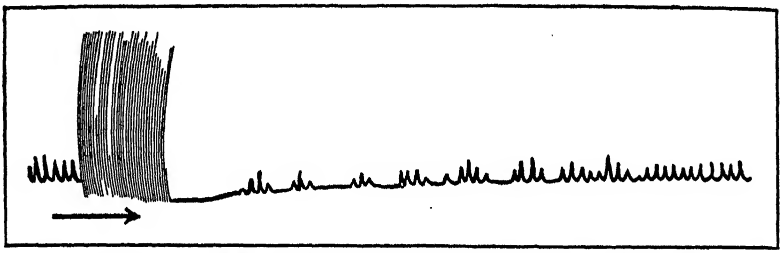


FIG. 247.—Periodic Breathing following Voluntary Hyperpnœa.

Record of respiration. After a period of voluntary hyperpnœa, apnœa develops, followed by periodic breathing.

ently of voluntary control. The bout of overventilation is followed by depressed breathing or by *apnœa* (cessation of rhythmic breathing). A phase of *periodic breathing* commonly sets in, after which respiration gradually becomes normal. Examination of the alveolar air at the end of the overventilation shows that the CO_2 tension is greatly reduced, *e.g.* to 15 mm. Hg (normal 40 mm.) and the O_2 tension is raised, *e.g.* to 140 mm. Hg (normal 100 mm.) (Figs. 247, 248).

The apnœa is due solely to the CO_2 lack (*i.e.* CO_2 tension below the minimum necessary for rhythmic breathing) and is not the result of the excess O_2 in the alveolar air. The following evidence may be quoted:

(i) If the overventilation is carried out with air containing 5% CO_2 so that the alveolar CO_2 concentration is not lowered, no apnœa occurs; in fact hyperpnœa persists after discontinuing the experiment although the alveolar O_2 concentration is raised to about 19%.

(ii) The inhalation of pure O_2 very markedly raises alveolar O_2 concentration, *e.g.* to 80%; but it does not alter alveolar CO_2 concentration and has no effect whatever on breathing.

During the period of apnœa, oxygen is being steadily removed from the alveoli, so that two minutes later, the alveolar oxygen tension is only 30 mm.

Hg; severe oxygen lack is thus present. At the same time CO_2 has been passing continuously from the tissues into the venous blood, diffusing out into the alveoli and remaining there. Thus the alveolar and arterial CO_2 tension slowly rises, reaching a level of about 36 mm. Hg (Fig. 248). Though this tension is still below the threshold level, "air hunger" is experienced and breathing is resumed. This resumption of the breathing is due to a combination of factors: (*a*) the rise of CO_2 tension almost to threshold, (*b*) the stimulating action of the severe anoxia. It can be shown that if either factor is lacking breathing does not recommence.

(*a*) An animal is vigorously overventilated for say 20 minutes. Alveolar O_2 tension rises to the same level as before (140 mm.) but alveolar, blood and

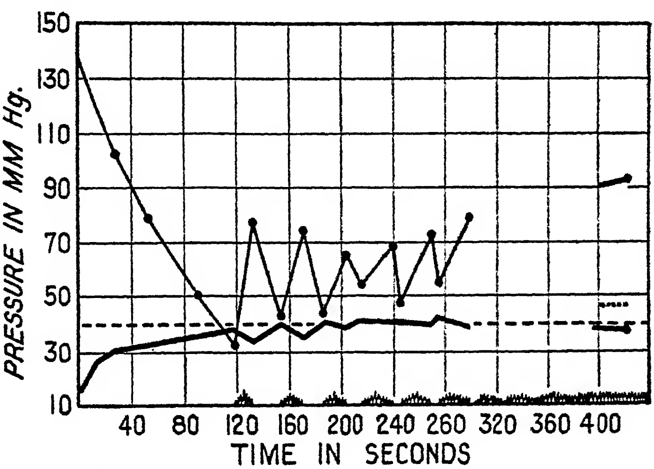


FIG. 248.—Changes in the Breathing and in the Tension of Gases in Alveolar Air after Forced Respiration for Two Minutes. (Douglas and Haldane.)

Upper curve, pressure of oxygen in alveolar air; Middle curve, pressure of CO_2 in alveolar air; Lower curve, respiration; Straight interrupted line = normal CO_2 tension (40 mm. Hg).

tissue CO_2 tensions are reduced to extremely low values. Apnœa sets in; after 2 or 3 minutes grave and progressive anoxia develops but breathing does not resume because the degree of CO_2 lack is too great to permit of anoxia alone acting as an effective stimulus to breathing. The animal may die of the anoxia without drawing a breath.

(*b*) If pure oxygen instead of air is taken in during the period of forced breathing, the apnœa is greatly prolonged, and in one man persisted for eight minutes. This is due to the fact that the resumption of the breathing has to await the rise of the CO_2 tension to its normal minimal value. Oxygen lack does not occur, and the oxygen tension is still high at the end of the period of apnœa. The breathing recommences gently without any sign of periodicity, breathing being maintained in a regular manner when CO_2 is the effective stimulus.

Under normal conditions *variations in the CO_2 tension alone* control breathing; but during the period of apnœa following voluntary overventilation,

the O_2 lack which develops enables breathing to be resumed before the CO_2 tension has risen fully to the normal threshold level. When breathing begins again, fresh oxygen supplies are taken in, the alveolar oxygen rises, and the oxygen want is temporarily relieved; at the same time, too, CO_2 is washed out from the alveoli and the arterial CO_2 tension falls. Thus both factors in the coalition of forces which brought about respiratory activity have been weakened. The breathing therefore ceases again. Similar events occur during the next period of apnœa: the alveolar oxygen is used up and increasing oxygen want develops; CO_2 accumulates and the alveolar and arterial CO_2 rise. Breathing resumes again. This cycle is repeated several times (Fig. 248). It may be wondered why the process does not continue indefinitely. The reason is that the brief bouts of breathing do not succeed in washing out CO_2 from the alveoli as rapidly as it is being turned out from the tissues and blood. The alveolar (and arterial) CO_2 tension is a little higher at the end of each period of apnœa than it was at the end of the preceding one, thus gradually rising to threshold level. Regular breathing is then finally established.

The results described are by no means invariably obtained. In some normal subjects apnœa does not follow the overventilation, but there is some decrease in the pulmonary ventilation; or the apnœa may pass smoothly into regular breathing of increasing amplitude. Sometimes the overventilation continues in spite of a marked fall of alveolar CO_2 tension and seems to be outside the subject's voluntary control.

The *general* effects of over-breathing are best studied when the pulmonary ventilation is increased to a more moderate extent (*e.g.* two- to threefold) and kept up for a longer time, *e.g.* 5 to as long as 30 minutes. Fig. 56 (p. 99) shows the resulting increased alkalinity of the blood, the lower CO_2 content of the plasma, and the usual compensatory changes to alkalæmia, namely, the excretion of a more alkaline urine with a lower NH_4^+ content (cf. p. 94); urinary volume and urinary phosphate are also increased. For some unknown reason *acetone bodies* appear in the urine, and the blood lactate is increased; these reactions may represent attempts to compensate for the alkalæmia by turning out fixed acids from the tissues into the blood. The pulse becomes rapid, sometimes irregular; the blood pressure is little altered; one would expect it to fall from the depressant effect of CO_2 lack on the vasomotor centre, but the *skin is pale, cold,* and moist; this cutaneous vasoconstriction which is due to the direct peripheral action of CO_2 lack may compensate for the vasodilatation of central origin (cf. p. 310). Leucocytosis and hyperglycæmia have been reported.

The *nervous* symptoms are numerous and important. There is a feeling of faintness, dizziness, unsteadiness, light-headedness, or of panic; the mind becomes more clouded and the surroundings seem unreal. There is numbness and tingling in the skin, especially in the hands, with impaired tactile sense. Typical symptoms of *tetany* develop (p. 1005), with carpo-pedal spasm and stiffness of the face and lips. Skilled voluntary movements are performed slowly and clumsily. The subject may show hysterical symptoms: he talks stupidly, laughs excessively, makes foolish grimaces, or sits with a vacant expression, the mouth open and the tongue protruding; consciousness may be lost. It is obvious that the *higher functions of the brain* are disturbed. Overventilation in man undoubtedly decreases the blood flow to the brain (p. 306); the resulting local anoxia may impair cerebral activity. There are

alterations in the electroencephalogram (p. 626) and the excitability of the motor cortex to electric stimulation is enhanced (p. 635).

Innervation of Bronchi.—The smooth muscle fibres of the bronchi receive dilator fibres from the sympathetic, and constrictor fibres from the vagi. Reflex stimulation of the latter fibres may be one of the causes of *asthma*. The afferent impulses producing this effect may arise from many parts of the body—*e.g.* from the nose (polypi), stomach, uterus, or from the bronchial tract itself. Histamine also induces intense bronchoconstriction (Fig. 205); it may possibly be liberated locally in various forms of bronchial infection (p. 338) and so set up spasm. In an asthmatic attack, inspiration, which is an active movement, can overcome the obstruction due to the diminished bronchial lumen, and succeeds in drawing air into the lungs. The passive return of the lungs to the position of rest which occurs normally in expiration is replaced by a laboured expiratory muscular effort. [For action of *adrenaline*, see p. 727.]

CARRIAGE OF OXYGEN IN THE BLOOD[1]

Certain essential data must be noted:

Alveolar Air: The oxygen content is 13–14%; the oxygen tension is about 100 mm. Hg (p. 364).

Arterial Blood—The average hæmoglobin content of blood is 15 g. per 100 c.c. As 1 g. of hæmoglobin when fully saturated combines with 1·34 c.c. of oxygen, the full oxygen combining power of blood is about 20 c.c. per 100 c.c. As arterial blood is only 95% saturated with oxygen, its oxygen content is about 19 c.c. [In this calculation the dissolved oxygen, about 0·3 c.c.% has, for simplicity, been ignored.] It must be emphasized that the arterial oxygen content will vary in different individuals according to their hæmoglobin content. The arterial oxygen tension is about the same as that in the alveolar air; in this book the alveolar and arterial oxygen tension is taken to be about 100 mm. Hg at sea level (cf. p. 367).

Mixed Venous Blood.—The oxygen content of the mixed venous blood depends on the degree of bodily activity. Under resting conditions it is about 14 c.c. O_2 per 100 c.c. blood, but in violent exercise it may be reduced to as little as 3 or 4 c.c. The oxygen tension correspondingly varies between 40 mm. Hg at rest and 20 mm., 15 mm., or less, during severe exertion.

The volume of O_2 which can be held in simple solution in the plasma at a tension of 100 mm. Hg is 0·3 c.c.%. As the actual volume of oxygen found in the arterial blood is about 19 c.c.%, it is clear that most of the gas is carried in loose combination with hæmoglobin.

Oxygen Dissociation Curves.—It is important to determine the amount of oxygen which hæmoglobin solutions or whole blood take up at different tensions of oxygen. To study this problem it is first necessary to expose a series of small samples of blood (*e.g.* 3 c.c.) in a glass vessel (of about 400 c.c. capacity) called a *tonometer* to various gas mixtures containing a known percentage of O_2 and therefore with a known O_2 tension. The bottle is rotated in a bath at a known temperature for half an hour. The blood

[1] Barcroft, *Respiratory Function of the Blood*, Cambridge, 1914, 2nd edn. pt. ii; *Hæmoglobin*, 1928; *Physiol. Rev.*, 1924, *4*, 329.

spreads out as a thin film over the sides of the bottle, comes into equilibrium with the gas and finally contains O_2 at the *same tension as that present in the tonometer*.

It is next necessary to determine (i) the *oxygen content—i.e.* how much oxygen the sample contains under the specified conditions; (ii) the *oxygen capacity—i.e.* the amount of oxygen the sample contains when it has been *exposed to air* (it is then fully saturated); (iii) the *percentage saturation—i.e.* the expression $\frac{\text{oxygen content}}{\text{oxygen capacity}} \times 100$.

DETERMINATION OF OXYGEN CONTENT OR CAPACITY.—The principle of the method is that potassium ferricyanide ($K_3Fe(CN)_6$) expels all the oxygen in oxyhæmoglobin (liberated from the cells by hæmolysis with saponin) with the formation of methæmoglobin.

To determine the O_2 *content*, the blood sample under investigation is introduced into the Van Slyke blood-gas apparatus; the red cells are laked by means of saponin and the contained oxyhæmoglobin is liberated. $K_3Fe(CN)_6$ is added; the volume of O_2 evolved is measured.

To determine the oxygen capacity another sample of blood is first vigorously *shaken in air* to saturate it with oxygen and then the amount of oxygen present is determined as before.[1]

If, for example, the oxygen content is 14 c.c. and the capacity 20 c.c. the percentage saturation is $\frac{14}{20} \times 100 = 70$.

The percentage oxygen saturation is determined of blood samples exposed to oxygen tensions from 0 to 100 mm. Hg pressure. The results are plotted with oxygen pressure as abscissa and percentage saturation as ordinate; in this way a *dissociation curve* is obtained which shows the relationship graphically. Curves are worked out for various conditions of temperature, H^+ ion concentration, hæmoglobin concentration, etc. Attention must again be drawn to the fact that 100% saturation represents an oxygen capacity which depends on the *amount of hæmoglobin* present—*e.g.* with 14 g. Hb%, the capacity is $14 \times 1{\cdot}34 = 18{\cdot}7$ c.c.%; with 15 g. Hb%, the capacity is $15 \times 1{\cdot}34 = 20$ c.c.%, etc. A consideration of the dissociation curves reveals many points of interest (Fig. 249).

At the oxygen tension of the arterial blood (about 100 mm. Hg) a *simple solution of hæmoglobin in water* becomes almost fully saturated with oxygen (about 95% saturated). At the oxygen tension in the mixed venous blood at rest, *e.g.* 40 mm. Hg, the hæmoglobin is still combined with most of its oxygen. In other words, if our blood consisted of a simple hæmoglobin solution, when it passed through the tissues very little oxygen would be evolved and death from anoxia would result. The curve for hæmoglobin is a rectangular hyperbola; that for *blood under the conditions found in the body* is **S**-shaped. Blood, too, is about 95% saturated at 100 mm. O_2 tension, taking up about 19 c.c.% of oxygen. But at a tension of 40 mm. it only carries about 14 c.c.% O_2, *i.e.* about 5 c.c.% have been made available for use by the tissues. It should be carefully noted that at tensions below 40 mm. the curve for blood descends steeply. This means that if the oxygen tension

[1] **Peters and Van Slyke, *Quantitative Clinical Chemistry Methods*, London, 1932, p. 257. For Haldane's blood gas analysis method, see Douglas and Priestley, *Human Physiology*, Oxford, 3rd edn., 1948.**

falls below 40 mm., *e.g.* to 20 or 15 mm. as in severe exercise, blood may give off 70 or 80% of its contained oxygen (*i.e.* about 14 or 16 c.c.%); greatly increased quantities of oxygen are thus supplied to the tissues (see continuous lines, Fig. 249).

The discrepancy between the behaviour of a simple hæmoglobin solution and whole blood is explained by the fact that blood in the body is at a temperature of 37° C., contains CO_2 at a tension of 40 mm. Hg (or more), and has a high concentration of hæmoglobin in the corpuscles. These facts can be readily demonstrated.

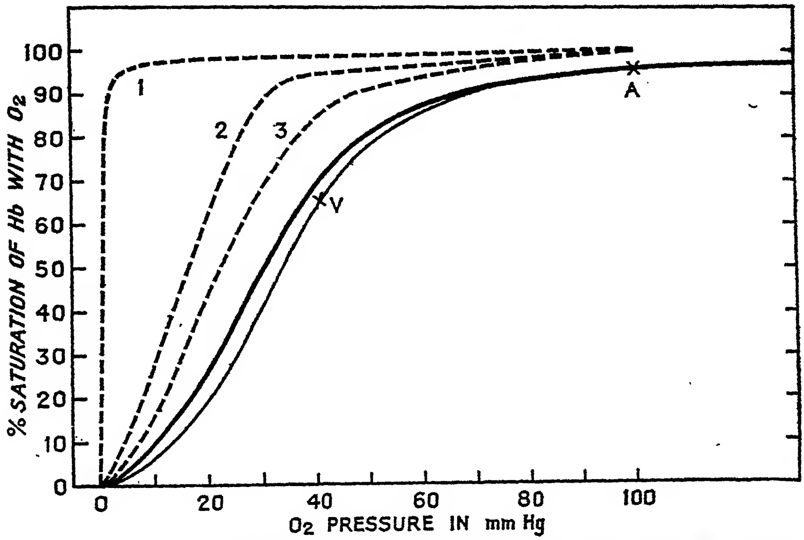


FIG. 249.—Diagram to illustrate the Behaviour of Solutions of Hæmoglobin and of Blood under Different Conditions.

Abscissa: Tension of Oxygen in mm. Hg. Ordinate: Percentage saturation of hæmoglobin with oxygen.
Dotted lines: Hæmoglobin solutions—1, Room temperature and no CO_2 tension; 2, Body temperature and no CO_2 tension; 3, Body temperature and 20 mm. CO_2 tension.
Continuous lines: Blood. Thick line—In presence of constant pressure of 40 mm. CO_2. Thin line—Dissociation curve in body when disappearance of oxygen is accompanied by increase of CO_2 pressure which facilitates dissociation of oxyhæmoglobin.
A = Arterial point; V = Venous point (mixed venous blood at rest).

(i) By raising the CO_2 tension to which a simple hæmoglobin solution is exposed, the dissociation curve is "shifted to the right," *i.e.* the convexity of the curve is flattened, which means that more oxygen is evolved at low oxygen tensions. CO_2 acts by altering the H^+ ion concentration of the solution. If the H^+ ion concentration is raised by any other acid, *e.g.* lactic acid, similar effects are produced.

(ii) Raising the temperature to that of the body, *i.e.* 37° C., has a similar though slight effect.

(iii) The strength of the hæmoglobin solution employed is of importance. A weak solution (*e.g.* 1 : 500) always gives a hyperbolic curve whether the reaction is acid, alkaline, or neutral. The strongest hæmoglobin solution that can be obtained gives an S-shaped curve under all circumstances. It is

noteworthy that the concentration of hæmoglobin reaches its maximum in the red blood corpuscles (about 30 g. of Hb per 100 c.c. of corpuscles), and that is presumably one of the chief factors producing the S-shaped dissociation curve of blood.

If the CO_2 tension, temperature, and hæmoglobin concentration are adjusted to that of the body, the solution behaves like blood.

Venous blood always contains CO_2 at a higher tension than arterial blood (46 mm. at rest and higher tensions during exertion). The dissociation curve for blood at *venous CO_2 pressure* is consequently shifted somewhat to the *right* compared with that for blood at arterial CO_2 pressure.

Rate of Oxygenation and Reduction of Hæmoglobin.—So far we have only considered the reactions of hæmoglobin when unlimited time is provided; but as the blood takes only 0·3–0·7 second to traverse a capillary in the lung the time available for oxygenation (and likewise for reduction in the tissues) is brief. The *velocity* of the hæmoglobin reactions must therefore be considered.

The apparatus of Hartridge and Roughton is employed. The essence of the method is that reduced blood is mixed with oxygenated Ringer's solution, and then caused to flow in a steady stream through a glass observation tube. The rate of the chemical reactions occurring in the tube can be measured with the reversion spectroscope which detects the slight alteration which takes place in the exact position of the absorption bands when reduced hæmoglobin is converted into oxyhæmoglobin. Combination with oxygen under these conditions is found to occur in *less than one-hundredth of a second*, *i.e.* more than fast enough. The reduction of oxyhæmoglobin also takes place with extreme rapidity and, according to Barcroft, it is speeded up by a rise of CO_2 tension or a rise of temperature.

To summarize: the *extent* of dissociation depends principally on (i) oxygen tension; (ii) CO_2 tension (and H^+ ion concentration); (iii) temperature; (iv) concentration of hæmoglobin.

The *rate* of dissociation depends on: (i) CO_2 tension; (ii) temperature.

Oxygen Carriage in the Body.—(1) Each 100 c.c. of arterial blood passes to the tissues carrying about 0·3 c.c. of oxygen in solution, and about 19 c.c. in combination with hæmoglobin. The tension of oxygen is about 100 mm. Hg; it must be remembered that tension is a property solely of the gas in *solution*. The oxygen tension in resting tissues is probably just a little lower than that found in the venous blood, *e.g.* about 35 mm. Hg. Owing to the great difference of oxygen pressure, oxygen rapidly passes out of the plasma through the capillary wall and tissue fluid to reach the tissue cells. The oxygen tension in the blood falls to about 40 mm. Hg; complete equilibrium with the tissues is not achieved. The oxyhæmoglobin in the corpuscles is now exposed to an O_2 tension of 40 mm. Hg. in the plasma around it, and therefore cannot retain all the oxygen it has hitherto held in combination. Dissociation occurs, and about 30% of the oxygen present (*e.g.* 5 c.c.%) is liberated. This volume of gas cannot remain in solution in the plasma, which already holds as much O_2 as it can; the oxygen liberated from the corpuscles must therefore diffuse out into the tissue fluid (Fig. 250).

As a result, the venous blood leaves with an oxygen tension of 40 mm. Hg and an oxygen content of about 14 c.c.%. There is slightly less oxygen in solution, and considerably less oxygen in combination with hæmoglobin.

The *arterio-venous oxygen difference* or difference in oxygen content of arterial and venous blood is 19—14=5 c.c.%.

In this instance the *coefficient of utilization*, which is

$$\frac{O_2 \text{ taken up by tissues}}{O_2 \text{ content of arterial blood}} = \frac{5}{19} = 0{\cdot}26, \text{ or } 26\%.$$

(2) When a tissue is *active, e.g.* the skeletal muscles during vigorous exercise, the venous blood becomes far more extensively reduced. Very active muscles may abstract almost all the oxygen brought to them in the arterial blood, *i.e.* the coefficient of utilization is extremely high. The mixed venous blood (that in the right auricle) may have an oxygen content of 7–8 c.c.% in hard work, or 3–4 c.c.% in extremely violent exercise, corresponding to coefficients of utilization of 65% and 80% respectively; as the mixed venous blood consists partly of blood from skin and viscera, which is only slightly reduced, it is clear that the blood from the active muscles must be almost completely free from oxygen. The essential factor concerned is the *fall of oxygen pressure in the blood* as it passes through the muscles; this

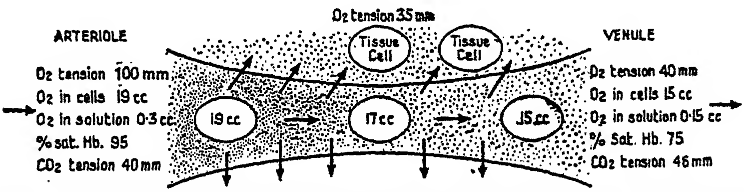


FIG. 250.—Diagram showing Passage of Oxygen from Blood to Tissues in Resting Organ.

causes almost complete dissociation of the oxyhæmoglobin. This more extensive reduction in active tissues is brought about as follows:

(i) The number of patent capillaries becomes greatly increased (*e.g.* 10 times), and all the capillaries become widely dilated (Fig. 191). Owing to the greatly increased total cross-section of the vascular bed locally, the *linear velocity* of the blood through the tissue is slowed down and thus more *time* is available for dissociation and diffusion of oxygen. The total blood flow is of course greatly increased because of local arteriolar dilatation (p. 432) and more general circulatory reactions (p. 433).

(ii) A large surface of blood is in contact at any one time with the tissues, and the gaseous interchange is further facilitated.

(iii) As the tissue is consuming O_2 at a great rate, the oxygen tension within it probably falls to zero. A very steep oxygen pressure gradient exists between plasma and tissues, permitting rapid diffusion.

The oxygen tension in the blood falls, *e.g.* to 30 mm. Hg; the oxyhæmoglobin dissociates and gives off about 60% of the oxygen combined with it. In an active tissue the temperature rises, larger amounts of CO_2 are evolved, and the H^+ ion concentration tends to go up. The rise of temperature and CO_2 tension increase the rate of dissociation of HbO_2. The increased CO_2 tension means that HbO_2 gives off *more oxygen at any given oxygen tension.* When exposed to a 30 mm. Hg O_2 tension, blood at 40 mm. CO_2 tension gives off half its combined oxygen; blood in which the decrease of O_2 is

accompanied by a corresponding increase of CO_2 pressure gives off 58% of its oxygen (Fig. 249). If the oxygen pressure falls to 10 mm., blood at a CO_2 pressure of 40 mm. becomes 10% saturated; but if the usual coincident rise of CO_2 pressure takes place the venous blood is actually only 7% saturated.

In the ways described activity of a tissue, by altering the calibre and number of the capillaries and changing the local O_2 and CO_2 tension and the temperature, may enable three times the volume of oxygen to be abstracted from the *same* volume of blood flow compared with the resting utilization. If the blood flow is also increased the total oxygen abstracted in unit time is correspondingly raised.

(3) When the mixed venous blood (O_2 tension 40 mm.) passes through the *lungs* it is exposed to an oxygen tension in the alveoli of 100 mm. Hg. Owing to the great difference of oxygen pressure, the gas rapidly diffuses from the alveoli through the thin pulmonary and capillary epithelium into the plasma (cf. p. 366). The amount of oxygen in solution is increased, and the tension in the plasma rises to about 100 mm. Hg. The hæmoglobin exposed to this high oxygen pressure takes up oxygen from the surrounding plasma and unites with it chemically. As oxygen is thus removed by the corpuscles from the plasma, fresh amounts of oxygen can pass from the lungs into the blood. A constant stream of oxygen thus passes from the lungs into the plasma and thence into combination with hæmoglobin. The arterial blood finally leaves the lungs almost fully saturated with oxygen (95% saturated), at an oxygen tension of 100 mm. Hg. The oxygen content is then about 19 c.c. (the exact figure depends on the hæmoglobin content); only 0·3 c.c. is in solution.

The peculiarities of *fœtal* hæmoglobin and the mechanism of *O_2 transport in the placenta between mother and fœtus* are discussed on pp. 1097 *et seq.*

CARRIAGE OF CO_2 IN THE BLOOD[1]

The following data about carbon dioxide in the blood must be remembered. They indicate the normal average findings in man. Variations above and below these figures are to be expected in healthy subjects:

TENSION OF CO_2: (i) in alveolar air, 40 mm. Hg (p. 364); (ii) in arterial blood, 40 mm. Hg (p. 367); (iii) in venous blood, 46 mm. Hg at rest (p. 279), and higher during exertion (up to 60 or 70 mm. Hg).

VOLUMES OF CO_2 PER 100 C.C. OF BLOOD: In *arterial blood* the normal range is 45–56 c.c.; in the example given below it was just over 48 c.c.; in *mixed venous blood at rest* it is 4 c.c. higher, *e.g.* 52 c.c.; in mixed venous blood during *severe exercise* a further marked increase occurs, *e.g.* up to 65 c.c.

The total CO_2 content of the plasma is nearly three times higher than that of the corpuscles.

CO_2 is present in the blood in three forms:

(i) In *simple solution* to form H_2CO_3.

(ii) As *bicarbonate,* mainly $NaHCO_3$ in the plasma and $KHCO_3$ in the corpuscles.

(iii) Combined with *hæmoglobin* to form *carbamino-hæmoglobin* in the

[1] See Peters and Van Slyke, *Quantitative Clinical Chemistry, Interpretations,* London, 1932. Haldane and Priestley, *Respiration,* new edn., Oxford, 1935. Roughton, *Physiol. Rev.,* 1935, *15,* 241.

corpuscles (and combined to a less extent with *plasma protein* to form a similar carbamino-protein).

(ii) and (iii) form the *combined* CO_2 of the blood.

Full representative data for arterial and mixed venous blood at rest are given in the Table below,[1] which should be carefully studied.

	Arterial Blood			Mixed Venous Blood		
	Plasma	Corpuscles	Whole Blood	Plasma	Corpuscles	Whole Blood
Volume in c.c.	60	40	100	59·6	40·4	100
% Saturation of Hb with O_2	..	95	..	..	75	..
pH	7·45	7·12	..	7·43	7·11	..
CO_2 pressure in mm. Hg	40	40	40	45·4	45·4	45·4
CO_2 in solution in c.c.	1·6	0·8	2·4	1·8	0·9	2·7
	Gain in dissolved CO_2 →			+0·2	+0·1	+0·3
Total combined CO_2 in c.c.	34·1	11·8	45·9	36·3	13·1	49·4
	Total Gain in Combined CO_2 →			+2·2	+1·3	+3·5
CO_2 as Bicarbonate	33·1	9·8	42·9	35·2	10·5	45·7
	Gain in CO_2 as Bicarbonate →			+2·1	+0·7	+2·8
CO_2 as Carbamino-compound	1·0	2·0	3·0	1·1	2·6	3·7
	Gain in CO_2 as Carbamino-compound →			+0·1	+0·6	+0·7
Total CO_2 in all forms	35·7	12·6	48·3	38·1	14·0	52·1
	Total Gain in CO_2 in all forms →			+2·4	+1·4	+3·8

Formation of Carbonic Acid. Carbonic Anhydrase.—The solution of CO_2 in water to form H_2CO_3, *i.e.* $CO_2+H_2O \rightarrow H_2CO_3$, takes place very slowly in pure water; the same is true of the reverse reaction, *i.e.* the evolution of CO_2 as a gas from its solution. An enzyme is present in the red blood corpuscles called *carbonic anhydrase* which enormously accelerates both the solution and the evolution of CO_2. The enzyme is destroyed by heat and inhibited by poisons like cyanide. The enzyme is not present in the plasma.

Carbon Dioxide Dissociation Curves.—Important information about the behaviour of the CO_2 of the blood under different conditions of CO_2 pressure (*e.g.* those found in the alveolar air or in the tissues) can be obtained by determining the CO_2 content of blood at different CO_2 pressures; the results, when plotted, constitute CO_2 dissociation curves; the abscissa represents *CO_2 pressure* in mm. Hg, the ordinate the *volume* of CO_2 in c.c. per 100 c.c. of blood. Curves are obtained for both *reduced* and *oxygenated* blood. The blood is exposed to an appropriate series of CO_2 pressures in a

[1] Stadie and O'Brien, *J. biol. Chem.*, 1937, *117*, 439

glass vessel (tonometer) until pressure equilibrium is reached; for reduced blood the diluent gas in the tonometer is hydrogen, for oxygenated blood a mixture of hydrogen and oxygen is used (the latter at a pressure of 100 mm. Hg which represents normal alveolar O_2 pressure). The blood sample is treated with acid (N/10 lactic acid) in a Van Slyke blood-gas analysis apparatus and exposed to a vacuum. The CO_2 of the blood, *in whatever forms it is present* (in combination or in solution) is evolved and the volume measured. Fig. 251 shows typical dissociation curves. As the CO_2 tension is increased the total CO_2 content rises (*i.e.* more CO_2 is taken up); as the CO_2 tension falls the CO_2 content falls (*i.e.* less CO_2 is taken up). In a vacuum the CO_2 content of blood is nil, *i.e.* all the contained CO_2 is given off completely. It will be noticed further that at any given CO_2 pressure, reduced blood takes up a larger volume of CO_2 than oxygenated blood; therefore under the conditions that exist in the body, reduction of the blood (*even without rise of CO_2 tension*) increases the CO_2 uptake by the blood; conversely, oxygenation of the blood leads to the evolution of CO_2. The straight line (dissolved CO_2) in Fig. 251 shows the volume of CO_2 in simple solution at different CO_2 pressures. The difference between these values and the *total* CO_2 content represents the CO_2 present in blood in combination. We must next consider in detail the combined forms of CO_2 in blood.

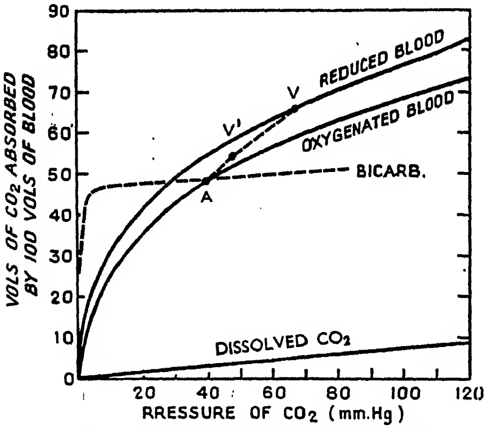


FIG. 251.—CO_2 Dissociation Curves of Blood.

Continuous curves: Lower curve = Dissociation curve of fully oxygenated blood; Upper curve = Dissociation curve of blood the hæmoglobin of which is fully reduced.
A = Arterial point (40 mm. Hg CO_2 pressure).
V = State of blood which has been fully reduced in body and in which each 19 c.c. of O_2 used has given rise to an additional 19 c.c. of CO_2 (R.Q. = 1).
AV = Dissociation curve of blood in body, when the increase of CO_2 content is accompanied by corresponding reduction of hæmoglobin.
V[1] = State of mixed venous blood at rest in the body (CO_2 pressure 45 mm. Hg).
Straight line: Dissolved CO_2 = CO_2 in solution.
Interrupted line: Bicarb. = Dissociation curve of bicarbonate solution. (Modified from Christiansen Douglas and Haldane.)

Carriage of CO_2 as Carbamino-Compounds.—There is satisfactory chemical evidence that CO_2 combines with the amino, *i.e.* the NH_2, groupings of hæmoglobin to form a compound called *carbamino-hæmoglobin* (in America it is known as *carbhæmoglobin*).[1]

$$Hb.NH_2 + CO_2 \rightleftarrows Hb.\ NH.COOH.$$

CO_2 combines in a similar manner though on a smaller scale quantitatively with the plasma proteins. The combination takes place directly between the *free* gas and the hæmoglobin (or plasma protein); it does not require previous

[1] This substance should not be confused with *carboxyhæmoglobin* (better called *carbon monoxy-hæmoglobin* or *carbonyl-hæmoglobin*), which is formed when carbon *monoxide* combines with hæmoglobin (p. 449).

solution of the CO_2 or the assistance of the enzyme carbonic anhydrase. Fig. 252 shows the relationship between CO_2 pressure and CO_2 uptake by hæmoglobin to form the carbamino-compound. At CO_2 pressures above 10 mm. Hg, hæmoglobin (either reduced or oxygenated) is almost completely saturated with CO_2 so that a rise of CO_2 pressure leads to little further CO_2 uptake *in this form*; the same holds good for the carbamino-protein. There is a striking difference, however, between the curves for reduced and oxygenated hæmoglobin; at any given CO_2 pressure reduced hæmoglobin takes up about 5 c.c. of CO_2 (per 100 c.c. of blood) more than does oxygenated hæmoglobin. The rise of CO_2 pressure in the tissues would thus not lead to any significant increase in CO_2 uptake in the carbamino-form; but the simultaneous reduction of oxyhæmoglobin which takes place considerably increases its affinity for CO_2 and the amount of CO_2 with which it combines (curve AV, Fig. 252).

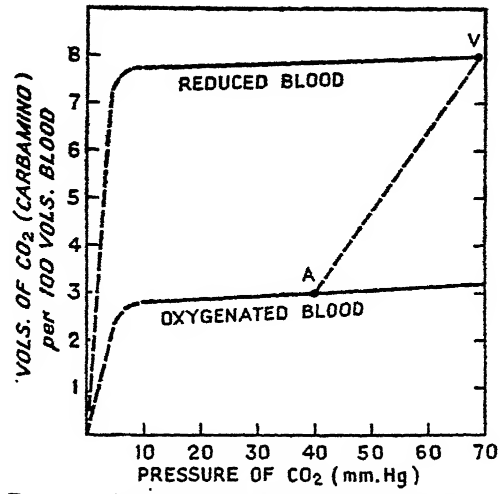


FIG. 252.—Carriage of CO_2 as Carbamino-compounds (mainly Carbamino-hæmoglobin).

Lower curve: oxygenated blood. Upper curve: reduced blood.
A: arterial point (oxygenated hæmoglobin, 40 mm. CO_2 pressure).
V: completely reduced venous blood under conditions in the body (fully reduced hæmoglobin, 70 mm. CO_2 pressure).
AV: CO_2 uptake as carbamino-compounds during passage of blood through tissues where rise of CO_2 pressure is associated with complete reduction of the hæmoglobin.

(Drawn by Prof. David Slome.)

The Table (p. 415) sets out these facts well; 100 c.c. of arterial blood contain in the carbamino-form 1·0 c.c. of CO_2 in the plasma (as carbamino-protein) and 2·0 c.c. in the corpuscles (as carbamino-hæmoglobin) at a CO_2 pressure of 40 mm. Hg. In the mixed venous blood at rest the amount in the plasma as carbamino-protein only rises to 1·1 c.c.; that in the corpuscles as carbamino-hæmoglobin rises to 2·6 c.c. owing to the simultaneous partial reduction of the hæmoglobin from 95% saturation to 75% saturation with O_2. When blood is fully reduced the amount of CO_2 as carbamino-hæmoglobin may rise to 7·0 c.c. (per 100 c.c. of blood).

Carriage of CO_2 as Bicarbonate.—The rest of the combined CO_2 of the blood is present, mainly in the plasma, and to a much less extent in the corpuscles, as bicarbonate. When the properties of bicarbonate in simple solution in water are examined they prove at first sight disappointing (Fig. 251, Bicarb.); bicarbonate is found to be a comparatively stable substance which (unlike blood) is entirely unaffected by variation of CO_2 pressure within the physiological range or even by low pressures of CO_2 down to a few mm. Hg. The only change produced is the expected rise or fall of the volume of CO_2 dissolved in the water. *In vacuo*, however, bicarbonate does dissociate, though

very slowly, giving off half its CO_2 content, the other half remaining combined with base as carbonate from which it can only be expelled by means of acid:

$$2NaHCO_3 \rightleftarrows CO_2 + Na_2CO_3 + H_2O.$$

In spite of these results, bicarbonate can be shown to play a very important part in CO_2 transport. It must be remembered that blood contains many substances which can modify the behaviour of the bicarbonate. It can be readily shown that the addition of hæmoglobin or plasma protein to a bicarbonate solution entirely alters its reactions to changes of CO_2 pressure; it now gives off or takes up CO_2 readily as the CO_2 pressure falls or rises; *in vacuo* it gives off its CO_2 rapidly and completely in the presence of hæmoglobin, and almost completely in the presence of plasma protein.

Mode of Action of Blood Proteins.—The mode of action of hæmoglobin and the plasma proteins can be explained very crudely as follows. Blood is an alkaline fluid containing basic and acid radicals. The base is first bound with the strong acid radicals, *e.g.* Cl, to form salts like NaCl. The base which is "left over" is shared between the two weak acids, namely, H_2CO_3 and the proteins (which act as weak acids in an alkaline medium), to form bicarbonate and proteinate respectively. Enough base is not available fully to neutralize these acids, and so we have left H_2CO_3, bicarbonate, "proteic acid" (protein acting as acid=HPr), and proteinate (protein combined with base=BPr.). There is constant competition between these two acids for base; the amount that each can obtain depends on the relative concentrations of proteic acid and H_2CO_3 in the blood. When the H_2CO_3 of the blood is increased, base is removed from the protein (using the term to include hæmoglobin and plasma protein) and more bicarbonate is formed, and *vice versa*.

$$\underset{}{H_2CO_3} + \underset{\substack{\text{(Protein}\\ +\text{base.)}}}{BPr} \quad \underset{\substack{\text{(Bicar-}\\ \text{bonate.)}}}{BHCO_3} + \underset{\substack{\text{(Protein}\\ -\text{base.)}}}{HPr}.$$

The equilibrium is thus a very unstable one, and is disturbed by variations in the mass of the interacting substances.

The protein chiefly under consideration is hæmoglobin, which is found in blood in both the reduced and the oxygenated form. The isoelectric point of reduced hæmoglobin is 6·8; that of oxyhæmoglobin is 6·65. At the isoelectric point a protein unites with neither acid nor base (p. 134). On the alkaline side of the isoelectric point proteins behave as weak acids; the farther away they are from their isoelectric point the stronger is their acid character. Their strength as acids can be measured by determining the amount of alkali they unite with. The results of such observations are shown in Fig. 253. It is seen that at any H^+ ion concentration oxyhæmoglobin unites with more base than does reduced hæmoglobin. At pH 7·25 (the normal reaction of blood) this difference can be expressed by representing oxyhæmoglobin as K_2HbO_2 and reduced hæmoglobin as $K_{1\cdot3}H_{0\cdot7}Hb$. (The base bound with hæmoglobin is potassium.) The oxidation of hæmoglobin causes it to unite with more base which is taken from the bicarbonate; CO_2 is consequently liberated and turned out from the blood. Conversely, reduction of the hæmoglobin causes it to give up base which is made available to unite with CO_2 to form bicarbonate. It follows that at any given CO_2

pressure, reduced blood will take up more CO_2 than does oxygenated blood. This amplifies from another standpoint the information already elicited from a study of the dissociation curves.

In vacuo, all the dissolved CO_2 is evolved and the concentration of H_2CO_3 falls to zero ; consequently the proteins of the blood (mainly the hæmoglobin) remove *all* the base from the bicarbonate, and all the CO_2 so combined is given off.

According to Van Slyke the changes undergone by hæmoglobin in the tissues can be represented thus :

$$(1)\quad K_2HbO_2 - O_2 + 0{\cdot}7H_2CO_3 \rightleftarrows K_{1{\cdot}3}H_{0{\cdot}7}Hb + 0{\cdot}7KHCO_3.$$

[In the lungs the reaction is :

$$(2)\quad K_{1{\cdot}3}H_{0{\cdot}7}Hb + O_2 + 0{\cdot}7KHCO_3 \rightleftarrows K_2HbO_2 + 0{\cdot}7H_2O + 0{\cdot}7CO_2.]$$

It is seen that 0·7 molecule of CO_2 can be taken up from the tissues and transformed into $KHCO_3$ for every 1 molecule O_2 used (equation (1)) *without any other assistance than the simultaneous reduction of hæmoglobin and its consequent weaker hold on base.* This bicarbonate is then broken up by the oxygenation of the hæmoglobin in the lungs (Equation (2)). If the R.Q. (CO_2 evolved/O_2 used) is 0·7 the above mechanism is adequate for CO_2 carriage. But if the R.Q. is 1 there is an additional 0·3 molecule CO_2 liberated (*i.e.* 1 molecule CO_2 for 1 molecule O_2). The extra CO_2 could, theoretically, be converted into bicarbonate by taking base from the hæmoglobin and plasma proteins as follows :

i. $H_2CO_3 + K.Hb \rightleftarrows KHCO_3 + HHb$.
ii. $H_2CO_3 + NaPr \rightleftarrows NaHCO_3 + HPr$.

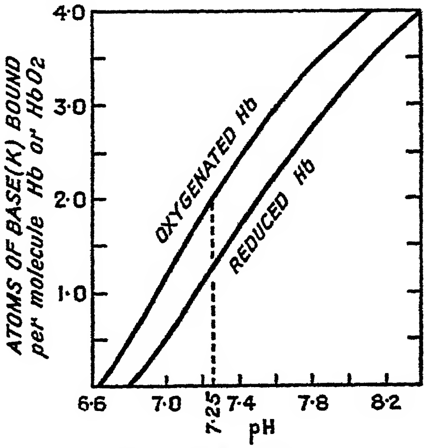


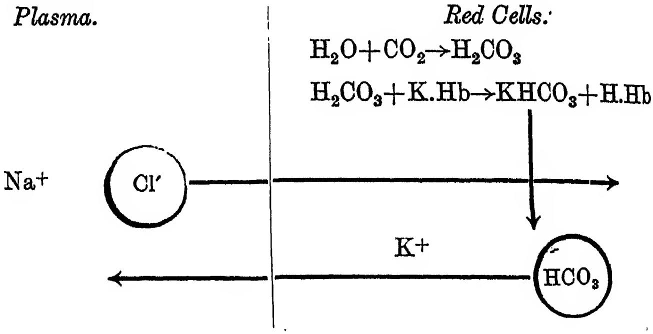
FIG. 253.—Base-binding Power of Oxy- and Reduced-Hæmoglobin.

The ordinate represents atoms of base (K) which unite with one molecule of hæmoglobin. Abscissa, reaction of medium.
At the isoelectric point no union with base occurs. The isoelectric point for Hb is 6·8, for HbO_2 it is 6·6.
At blood pH (7·25) each molecule of HbO_2 units with 2 atoms of K, and each molecule of Hb with 1·3 atoms. Reduction of oxyhæmoglobin liberates 0·7 atom of K per molecule. This is available for union with CO_2 to form bicarbonate. If more base is then taken from reduced hæmoglobin, the pH moves to the left, *i.e.* the reaction becomes more acid. (Peters and Van Slyke, *Quantitative Clinical Chemistry, Interpretations*, 1932.)

In view of the *enhanced affinity of reduced hæmoglobin for CO_2 to form carbamino-hæmoglobin*, it is unlikely that this series of reactions is called upon to any significant degree.

INTERCHANGES BETWEEN THE PLASMA AND CORPUSCLES.—As explained on p. 7 the red cell membrane is permeable to the negatively charged anions (*e.g.* Cl′, HCO_3') and to H^+ ions; it behaves, however, as though it were impermeable to the other positively charged cations (*e.g.* Na^+, K^+). When CO_2 enters the blood from the tissues there is only a small

increase in the H_2CO_3 content of the plasma (owing to absence of carbonic anhydrase), but there is a large increase of H_2CO_3 in the corpuscles (owing to the presence of anhydrase). H_2CO_3 then reacts with hæmoglobin to form much additional bicarbonate ; very little bicarbonate is formed primarily in the plasma. As the HCO_3' concentration of the cells is raised it migrates out into the plasma ; an equivalent amount of Cl′ moves from the plasma into the cell. These reactions can be shown thus :



These migrations can be proved experimentally. If the CO_2 tension of whole blood is raised the chloride of the plasma decreases, the chloride of the corpuscle increases and the plasma bicarbonate increases. This is known as the *chloride shift* or *Hamburger phenomenon* ; it might be better called the *bicarbonate shift,* so drawing attention to the physiologically significant aspect of the exchange. Without this ionic exchange almost the entire uptake of CO_2 as bicarbonate would occur in the red cells ; because of this exchange most of the CO_2 *primarily converted into bicarbonate in the cells* is transferred secondarily to the plasma to form $NaHCO_3$.

The Table on p. 415 shows these facts quantitatively. In the example given, 100 c.c. of arterial blood contained 33·1 and 9·8 c.c. of CO_2 as bicarbonate in plasma and corpuscles respectively ; in the mixed venous blood the corresponding values were 35·2 and 10·5 c.c. The plasma has thus gained (chiefly through the ionic shift just described) 2·1 c.c., and the corpuscles only 0·7 c.c. as bicarbonate. The total uptake as bicarbonate was 2·8 c.c.

When CO_2 is taken up more H_2CO_3 is formed, and the H^+ ion concentration tends to rise ; but this is largely compensated by the simultaneous formation of bicarbonate which gives rise to an excess of OH′ ions ; the reverse changes take place in the lungs. The various reactions described above buffer the blood so effectively that the difference in *p*H between arterial and venous blood at rest is 0·01–0·02. As explained above, bicarbonate is readily formed in the corpuscles and thus their reaction is well protected (*primary buffering*). The comparatively small direct formation of bicarbonate in the plasma means that primarily it is poorly buffered. The outflow of HCO_3' from the corpuscles into the plasma increases the bicarbonate content of the plasma and so preserves its normal reaction (*secondary buffering*).

The bicarbonate of the plasma has three functions : (i) it acts as a means of *CO_2 transport* ; (ii) it is the *alkali* of the plasma (the reaction of plasma

depends on the ratio of H_2CO_3 to $BHCO_3$); (iii) it is the *alkali reserve* which neutralizes any strong acids which enter the blood (p. 92).

CO_2 Dissociation Curve in the Body.—The dissociation curves in Fig. 251 for reduced and oxygenated blood represent somewhat artificial conditions. In the body the uptake of CO_2 by blood as it passes through the tissues is accompanied by a simultaneous and corresponding evolution of O_2; similarly the evolution of CO_2 in the lungs is accompanied by oxygenation of the blood. When 100 c.c. of blood are fully reduced in the tissues 19 c.c. of O_2 are evolved. The amount of CO_2 formed varies with the local respiratory quotient. Taking the two extreme values for the R.Q. of 0·7 and 1·0, the CO_2 formed is 13·3 c.c. and 19·0 c.c. respectively; for intermediate values of the R.Q. intermediate volumes of CO_2 are formed. Thus with an R.Q. of 0·85 an additional 16·0 c.c. of CO_2 are evolved. The CO_2 content of fully reduced blood *in the body* (when the arterial CO_2 content was 48 c.c.%) would vary between $48+13·3=61·3$ c.c. at R.Q. 0·7, and $48+19=67$ c.c. at R.Q. 1·0; at R.Q. 0·85 it would be $48+16=64$ c.c. Inspection of Fig. 254 shows that fully reduced blood with a CO_2 content of 67 c.c. has a CO_2 tension of 70 mm. Hg. This point is marked off on the upper (reduced blood) curve and labelled V. The point A on the (lower) oxygenated blood curve represents the condition in the arterial blood. The line drawn from A to V represents the CO_2 dissociation curve of blood in the body (when the R.Q. is 1), and the rising CO_2 pressure and CO_2 uptake are associated with a falling O_2 pressure and progressive reduction of the hæmoglobin. The point V^1 represents mixed venous blood at rest (CO_2 pressure at 45 mm. Hg + an O_2 saturation of hæmoglobin of 75%).

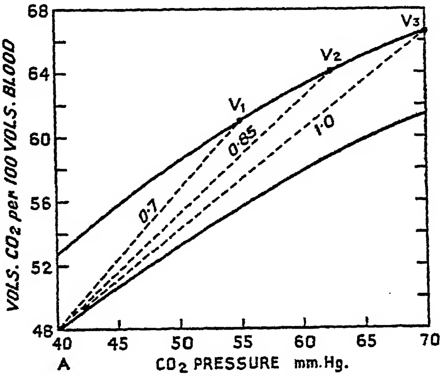


FIG. 254.—CO_2 Dissociation Curves in the Body at Varying Values for the Respiratory Quotient. (Modified from Douglas and Priestley.)

This Figure is an enlarged and amplified reproduction of the central portion of Fig. 251 (*q.v.*).
Upper curve: dissociation curve of blood the hæmoglobin of which is fully reduced.
Lower curve: dissociation curve of fully oxygenated blood (19 c.c. of O_2 per 100 c.c. blood).
The arterial point A, is at a CO_2 pressure of 40 mm. Hg on the lower curve.
V^3: state of venous blood which has been fully reduced in the body at R.Q.=1. (CO_2 content is $48+19=67$ c.c.; CO_2 pressure, 70 mm. Hg.)
V^2: the same, at R.Q. 0·85 (CO_2 content is $48+16=64$ c.c.; CO_2 pressure, 63 mm. Hg.).
V^1: the same, at R.Q. 0·7 (CO_2 content, $48+13·3=61·3$ c.c.; CO_2 pressure, 55 mm. Hg.)
(The line AV^3 corresponds to the line AV in Fig. 251.)

Fig. 254 shows the dissociation curves in the body when the R.Q. is 0·7 (curve AV^1), when the R.Q. is 0·85 (curve AV^2) and when the R.Q. is 1·0 (curve AV^3). The lower the R.Q., the smaller is the additional CO_2 uptake and the lower is the CO_2 tension in completely reduced venous blood.

Fig. 255 (which should be studied in conjunction with the discussion of CO_2 transport (p. 422) shows the various forms in which the *extra* CO_2 is taken up by the blood when it passes from the arterial point to the completely

reduced venous point at an R.Q. of 0·85 (corresponding to the curve AV^2 in Fig. 254), *i.e.* when the rise of CO_2 tension is associated with an equivalent reduction (deoxygenation) of the hæmoglobin.

CO_2 Transport in the Body.—(1) INTERCHANGES IN THE TISSUES.—(i) Arterial blood reaches the tissues with a CO_2 tension of 40 mm. Hg and a CO_2 content of 48 c.c.%. The CO_2 tension in resting tissues is higher (about 46 mm. Hg); CO_2 in solution diffuses rapidly into the blood and the CO_2 tension in the venous blood rises (at rest) to 46 mm. Hg; the CO_2 content rises to 52 c.c.%.

(ii) The reaction $CO_2 + H_2O \rightarrow H_2CO_3$ takes place slowly in the plasma but very rapidly in the corpuscles owing to the presence there of carbonic

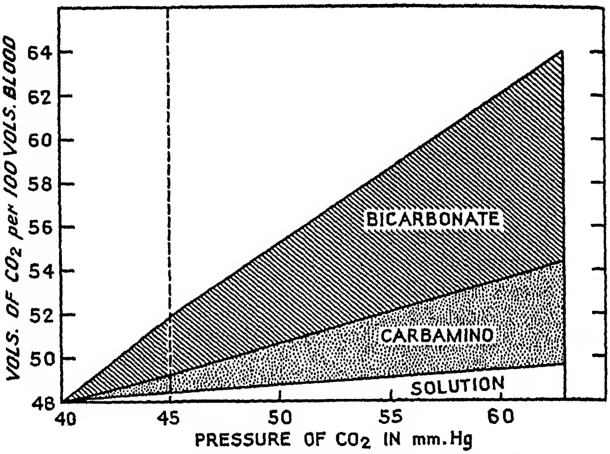


FIG. 255.—Uptake of CO_2 in Various Forms when the rise of CO_2 Tension is associated with a corresponding De-oxygenation of the Hæmoglobin (R.Q.=0·85).

Arterial Point: 40 mm. Hg CO_2 pressure; CO_2 content, 48 c.c.%.
Mixed Venous Blood at rest: 45 mm. Hg CO_2 pressure; CO_2 content, 51·8 c.c.%.
Fully Reduced Venous Blood (R.Q.=0·85): 63 mm. Hg pressure; CO_2 content, 64 c.c.%. (Drawn by Prof. David Slome.)

anhydrase. The amount of H_2CO_3 in solution (mainly in the corpuscles) is increased, and the equilibrium will be disturbed in its favour (Fig. 256).

(iii) Base passes mainly from the hæmoglobin and slightly from the plasma protein; fresh bicarbonate is formed. Simultaneously, however, oxyhæmoglobin has yielded up some oxygen and has become partially reduced. It has thus become a weaker acid with a feebler hold on base; the H_2CO_3 is stronger. Base is therefore taken, mainly from hæmoglobin, to form bicarbonate (Fig. 253). Most of the bicarbonate primarily formed in the corpuscles is secondarily transferred to the plasma (Hamburger shift).

(iv) At the same time CO_2 combines directly with the partially reduced hæmoglobin to form carbamino-hæmoglobin (and also with plasma protein) (Figs. 252, 255).

(v) As CO_2 is removed from solution in these various ways, room is made for more CO_2 to pass into the blood from the tissues. Venous blood at rest takes up about 4 c.c.% of CO_2. The Table on p. 415 shows in detail

the differences between arterial and resting venous blood. The reactions involved are summarized in Fig. 256 and the Table below.

(vi) During *exertion* and in tissue activity generally, the local CO_2 tension is higher (*e.g.* 60–65 mm. Hg) and the venous CO_2 tension is correspondingly raised. Simultaneously the oxyhæmoglobin is more extensively reduced. Larger amounts of base are transferred, mainly from hæmoglobin, and a great increase in the bicarbonate formed occurs; at the same time more carbamino-hæmoglobin is formed. The venous blood may contain about 64 c.c.% CO_2 (chiefly in combination) under these conditions (cf. Table below and Fig. 256 for full details). The dilatation and increase in number of the capillaries and the other circulatory changes both local and general, help the taking up of CO_2 in the same way as they aid the giving off of oxygen (cf. p. 413).

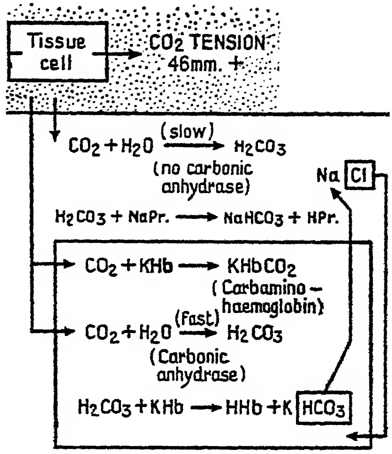


FIG. 256.—Diagrammatic Representation of CO_2 Uptake by Blood from Resting Tissues.

(2) INTERCHANGES IN THE LUNGS.—(i) In the lungs, the venous blood (CO_2 tension at rest = 46 mm.) is exposed to an alveolar CO_2 pressure of 40 mm. Hg. Owing to the difference of pressure on the two sides of the membrane, CO_2 passes out of the blood into the lungs. The preliminary reaction $H_2CO_3 \rightarrow H_2O + CO_2$ is catalysed by carbonic anhydrase in the corpuscles.

	Arterial Blood.	Resting Venous Blood.	Fully reduced Venous Blood. [R.Q.=0·85]
CO_2 pressure	40 mm.	45·4 mm.	63 mm.
O_2 content %	19·0 c.c.	15·0 c.c.	0 c.c.
% Saturation of Hb	95	75	0
CO_2 in solution %	2·4 c.c.	2·7 c.c. [+0·3]	3·9 c.c. [+1·5]
Total CO_2 % (all forms)	48·3 c.c.	52·1 c.c. [+3·8]	64·4 c.c. [+ 1·5 in solution + 14·6 in combination]
CO_2 as carbamino-compounds %	3·0 c.c.	3·7 c.c. [+0·7]	8·0 c.c. [+5·0]
CO_2 as Bicarbonate %	42·9 c.c.	45·7 c.c. [+2·8]	52·5 c.c. [+9·6]

As previously pointed out (p. 367), CO_2 diffuses very rapidly through the pulmonary epithelium, so that a pressure difference of 6 mm. Hg is quite sufficient to enable the appropriate amounts of CO_2 to be evolved. The CO_2 pressure in the alveoli does not rise in spite of the addition of CO_2, which is being turned out from the blood, because the breathing is regulated at a level which is

sufficient to wash out as much CO_2 as is coming out from the blood. The CO_2 tension in the blood, when arterialized, falls to 40 mm. Hg, *i.e.* pressure equilibrium with the alveolar air is attained. The amount of H_2CO_3 in solution in the blood is therefore diminished and its "mass action" is reduced.

(ii) The blood has meanwhile become oxygenated; oxyhæmoglobin is formed, which is a stronger acid with a greater avidity for base than reduced hæmoglobin. The equilibrium is thus disturbed in favour of the proteins of the blood. They take base from the bicarbonate and form more hæmoglobinate (and, slightly, proteinate). The CO_2 which has been deprived of its base cannot remain in the blood, which already holds as much CO_2 as it can in solution, and consequently it must pass out into the lungs.

(iii) At the same time carbamino-hæmoglobin releases its CO_2 owing to the simultaneous oxygenation of the hæmoglobin.

(iv) As the bicarbonate of the corpuscles is disintegrated and CO_2 is evolved, the Hamburger shift is reversed. Bicarbonate migrates from the plasma to the corpuscles; H_2CO_3 is formed which gives off its CO_2 owing to the action of carbonic anhydrase; chloride returns from the corpuscles to the plasma.

(v) Very little *direct* evolution of CO_2 from bicarbonate takes place in the plasma for two reasons:

(*a*) owing to the absence of a potent competing protein;

(*b*) any H_2CO_3 formed is further handicapped in giving off its CO_2 because of lack of carbonic anhydrase. The liberation of the combined CO_2 of the plasma thus takes place in the following stages:

(*a*) migration of HCO_3' from the plasma into the corpuscles to unite with the available K^+ (freed by Cl' shift) to form $KHCO_3$;

(*b*) interaction of $KHCO_3$ with oxyhæmoglobin to liberate H_2CO_3;

(*c*) liberation of CO_2 from H_2CO_3 by carbonic anhydrase;

(*d*) passage of free CO_2 from the corpuscles back into the plasma and thence into the alveoli of the lungs. About 4 c.c. of CO_2 are given off in the lungs by 100 c.c. of blood at rest and more during activity.

Inter-relation of O_2 and CO_2 Transport.—Reduction of the blood increases CO_2 uptake; conversely, CO_2 uptake helps the expulsion of O_2 from the blood, *i.e.* it shifts the O_2 dissociation curve to the right. Oxygenation of the blood increases expulsion of CO_2: loss of CO_2 from the blood has, however, little effect on O_2 uptake in the lungs.

Hæmoglobin=Hæm+Globin.

Hæm may be represented as Por: Fe^{++} (pp. 93, 174); globin as $R\langle{}^{NH_2}_{COOH}$ (to indicate the presence of amino and carboxyl groupings.[1]

Oxyhæmoglobin = Por: $Fe^{++}O_2\left[R\langle{}^{NH_2}_{COOK}\right]$, *i.e.* it is combined with O_2 and with base (K).

Reduced hæmoglobin = Por: $Fe^{++}\left[R\langle{}^{NH.COOH}_{COOH}\right]$, *i.e.* it has gained CO_2 to form a carbamino-compound and has lost both O_2 and base (K).

[1] Note in the equations that *oxygen* combines with *hæm*, while CO_2 is carried by the *globin* part of the hæmoglobin molecule.

The changes in the tissues can be shown thus:

$$CO_2 + Por : Fe^{++}O_2\left[R\begin{matrix}NH_2\\COOK\end{matrix}\right] \longrightarrow Por : Fe^{++}\left[R\begin{matrix}NH.COOH\\COOH\end{matrix}\right] + O_2 + KHCO_3.$$

Decompression Sickness.[1]—This term is applied to the syndrome which is produced when the external pressure is rapidly and markedly decreased. The condition may occur in two sets of circumstances: (i) when a high external pressure (several atmospheres) is suddenly lowered to normal; (ii) when a normal external pressure is suddenly reduced to a very low level.

(1) The first group to be considered includes deep divers and workers in caissons (steel chambers which are sunk in water and filled with compressed air), who are exposed to the effects of raised atmospheric pressure. At this stage there is little discomfort except temporary giddiness and noises in the ears. The breathing and pulse are slowed and there is frequency of micturition. Symptoms appear when rapid *decompression* takes place, *i.e.* when the subject returns to normal atmospheric conditions; they may set in immediately or after some hours. The common symptoms are: pain in the limbs which cause the victim to double up in anguish (the "bends") and a choking sensation (the "chokes"). In severe cases there is extensive paralysis, *e.g.* from the waist downwards. There may be signs of circulatory failure and cyanosis; in very bad cases death may occur in a few minutes.

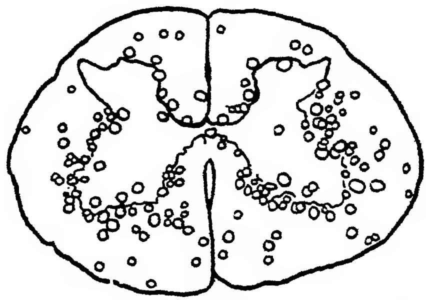


FIG. 257.—Nervous System in Decompression Sickness (Compressed Air Sickness, Caisson Disease).

This shows the distribution of the gas bubbles which are formed in the spinal cord during decompression. (Haldane and Priestley, *Respiration*, 1935.)

These symptoms can be readily accounted for. Under high atmospheric pressure, excess nitrogen (and oxygen) are taken up in solution. On rapidly reducing the pressure, nitrogen is evolved from solution and aggregates into bubbles; the mechanism of bubble formation is obscure. The volume of gas evolved in any region depends on the solubility of the gas in the tissue. Nitrogen is about five times as soluble in fats (*e.g.* adipose tissue and lipoids of the central nervous system and peripheral nerves), as in water. Wherever bubbles appear they tear the local tissue, press on or damage nerves, or block small blood vessels. The bends are probably due to damage to tissues surrounding the joints. In severe cases, bubbles form in the central nervous system, especially in the spinal cord at the junction of the white and grey matter (Fig. 257); it is thought that general circulatory failure is caused by bubbles accumulating in the blood in the heart, the resulting frothing preventing the expulsion of the blood from the ventricles into the great vessels. Bubbles may also appear in the cerebrospinal fluid, the pressure of which may rise by some 3 cm. H_2O. The excess oxygen which is evolved is used by the tissues and causes no trouble.

[1] Catchpole and Gersh, *Physiol. Rev.*, 1947, *27*, 360. Fulton, *Aviation Medicine in its Preventive Aspects*, London, 1948.

To prevent the onset of the symptoms the pressure in the caisson must be reduced in stages to enable the dissolved nitrogen to be gradually given off. If symptoms appear, recompression must be carried out at once.

The syndrome described above, as it occurs in caisson workers, is often called *caisson disease* or *compressed air sickness,* although, as already explained it is not the compression but the decompression which causes the trouble.

(2) The second group includes :

(i) Aviators who ascend very rapidly from sea level to a great height; in consequence the nitrogen normally present in the body fluids is evolved from solution. If they breathe oxygen for 3–4 hours before the ascent, they can eliminate most of this nitrogen via the lungs. They can subsequently ascend at a rate of 5000 feet per minute to 37,000 feet without showing symptoms of decompression sickness.

(ii) Subjects who are experimentally exposed to a low pressure in a decompression chamber.

(iii) People flying at great heights in aircraft with pressurized cabins; if the cabin bursts they are instantaneously exposed to extremely low pressures (" explosive decompression "). In this group the symptoms are due to anoxia combined with the effects of decompression.

Nitrogen Poisoning.—When air is breathed at 8·5–10 atmospheres pressure, signs of mental deterioration set in rapidly owing to the narcotic action of nitrogen at these pressures; the symptoms are not produced by mixtures of oxygen+helium at the same pressure.

SKELETAL MUSCLE.[1] MUSCULAR EXERCISE

Constituents of Muscle.—The principal constituents of skeletal muscle are:

(i) *Water*, about 80%.

(ii) *Protein* (17%), chiefly *myosin* and *actin*; these two proteins acting together (*acto-myosin*) are believed to be the contractile elements, the change in the configuration or in the arrangement of the molecules being responsible for muscular contraction or relaxation. Associated with myosin is an enzyme, adenosine triphosphatase (*ATP-ase*) which acts on adenosine *tri*phosphate (ATP) breaking it down to adenosine *di*phosphate (ADP). There are many other protein enzymes and their organic (non-protein) coenzymes in muscle (as in all other tissues) which are responsible for the many complex chemical transformations that take place there.

(iii) *Glycogen* (in concentrations up to 1%) represents a readily available reserve of energy that is called upon during muscular activity. It cannot be directly transformed into blood sugar and is not utilizable for the relief of hypoglycæmia. During its breakdown (" dissimilation ") it gives rise to numerous intermediaries, many of them combined with phosphate, as detailed on p. 846. Muscle can also utilize the *ketone bodies* reaching it in the blood (from the liver) as a source of energy. Other *undetermined substrates* may also be used.

[1] Barer (Structure of Muscle fibre), *Biol. Rev.*, 1948, *23*, 159. Symposium on Muscular Contraction, *Ann. N.Y. Acad. Sci.*, 1947, *47*, 665–930 (including Dynamics, Ultrastructure, Chemistry, Mechano-chemical Coupling). Mommaerts, *Muscular Contraction*, N.Y., 1950. Szent-Györgyi, *Chemistry of Muscular Contraction*, 2nd edn., N.Y., 1951.

(iv) *Adenosine triphosphate* (ATP): as illustrated in Fig. 553, two of the three phosphate bonds are high-energy phosphate bonds (indicated by the symbol ~ph). ATP is the *immediate* source of energy of muscular activity; one of its two high-energy phosphate bonds is rapidly split off by the enzyme *ATP-ase* leaving ADP.

(v) *Creatine phosphate* (average concentration 0·5%): its one phosphate group also carries a high-energy bond; this substance thus serves as a convenient *storehouse* of energy which is called upon when necessary to reform ATP from ADP (p. 843).

(vi) There are a number of specific muscle constituents to which no function has as yet been assigned; a striking example is *carnosine* (β-alanyl-histidine).

Cause of Muscular Contraction and Relaxation.[1]—The arrival of a nerve impulse at the motor end-plate releases acetylcholine (p. 513) which

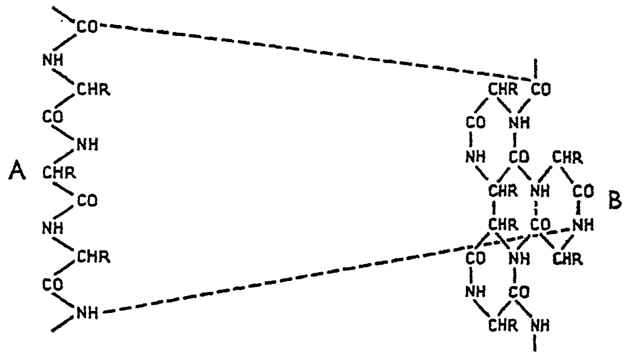


FIG. 258.—Relaxed and Contracted Form of Myosin Molecule. (Modified from Astbury, *Ann. Rev. Biochem.*, 1939, *8*, 113.)
Left (A): in fully relaxed condition.
Right (B): in contracted (shortened) state.
R=Amino-acid side chains.

in turn sets up the end-plate potential (p. 512). When this attains a critical magnitude it sets up in the muscle fibre a propagated action potential which is followed by mechanical and chemical changes. Muscular contraction may be isometric (development of tension without shortening) or isotonic (development of tension plus shortening); contraction is followed by relaxation. The mechanical changes are due primarily to alterations in the state of the two proteins, myosin and actin (acto-myosin). Thus acto-myosin has been extracted from muscle in the form of threads which *in vitro* can be made to contract by adding ATP to a medium containing a suitable concentration of Mg^{++} or K^{+} ions. Contraction may be due to an alteration in the configuration of the protein molecules, *e.g.* the amino-acid groups may become folded up on themselves as shown in Fig. 258; alternatively, it has been suggested that the protein molecules are dispersed in the relaxed state and become closely packed together and interlocked during contraction.

Two views have been expressed about the nature of muscular activity:

(i) That muscular relaxation is the active phase of the cycle. The state

[1] Symposium, Hill *et al.*, *Proc. roy. Soc. B.*, 1950, *137*, 40. Hill, *ibid.*, 268.

of the relaxed muscle is likened to a pulled out taut spring; energy must be applied to produce and maintain this condition. Contraction is attributed to an undefined agent which permits the "spring" to recoil. The potential energy which was stored by the muscle during "relaxation" is expended during the succeeding contraction in varying proportions as work and heat. On this hypothesis heat could be liberated during and after contraction but not before it.

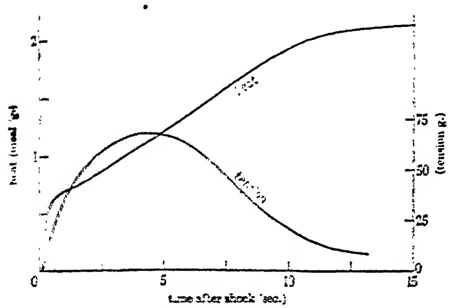


Fig. 258A.—Time Relation between Heat Production and Development of Tension. (Hill, *Proc. roy. Soc. B.*, 1950, *137*, 269.)

Tortoise muscle at 0° C. Abscissa: time after stimulus—shock in seconds. Ordinate: left, heat evolved in millicalories per g.; right, tension in g.

(ii) That muscular contraction is the active phase of the cycle. On this hypothesis some chemical changes occur first which release energy that is used in causing development of tension or shortening. If this view is correct heat might be evolved in stimulated muscle before the rise of tension.

The decisive experiment consists in determining the exact time relationship between the first appearance of heat and of tension in stimulated muscle. Observations have been made on stimulated tortoise muscle cooled to 0° C. in which the course of events is greatly slowed down and thus more readily followed. It is found that the evolution of heat begins before the mechanical response indicating that muscular contraction is the active phase of the cycle (Fig. 258A).

Chemical and Energy Changes associated with Muscular Activity.—These are worth studying for their own sake, though as has been emphasized their relationship to muscular activity is obscure.[1] Some of the reactions to be described occur as readily in the absence of oxygen (*anærobically*) as they do in the presence of oxygen (*ærobically*); others can only take place in the presence of oxygen.

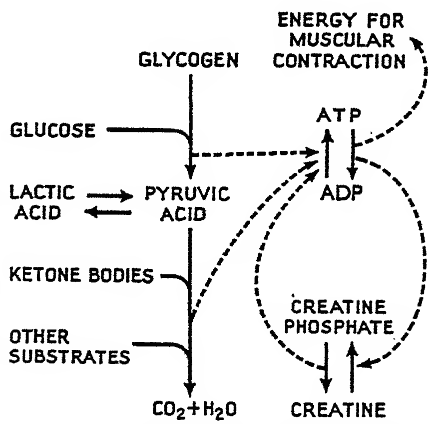


Fig. 259.—Summary of Chemical Changes in Muscular Activity.

ADP, ATP, Adenosine diphosphate, adenosine triphosphate.

[1] Hill, "A Challenge to Biochemists," *Biochem. Biophys. Acta*, 1950, *4*, 4.

The following are the principal known chemical reactions associated with muscular activity (Fig. 259):[1]

(1) ADENOSINE TRIPHOSPHATE (ATP)→ADENOSINE DIPHOSPHATE (ADP).—The initial event is the breakdown of adenosine triphosphate (ATP) to the diphosphate (ADP). The specific enzyme involved is *adenosine triphosphatase* (*ATP-ase*) which is found in association with myosin; high-energy phosphate bonds (∼ph) are released (p. 843).

$$ATP \rightarrow ADP + {\sim}ph.$$

The above reaction provides "bond energy" which is in some unexplained way responsible for contraction. This reaction can occur anærobically.

(2) BREAKDOWN OF CREATINE PHOSPHATE.—The next event is the breakdown of creatine phosphate to yield creatine and a high-energy phosphate bond. The latter is applied to ADP, steadily rebuilding it into the higher energy state of ATP, thus:

$$\text{Creatine Phosphate} \rightarrow \text{Creatine} + {\sim}ph.$$
$${\sim}ph + ADP \rightarrow ATP$$

This reaction can occur anærobically (p. 893).

As a result of reactions (1) and (2) adenosine triphosphate is always available to supply energy for muscular activity until all the creatine phosphate has been finally broken down to creatine.

(3) BREAKDOWN OF MUSCLE GLYCOGEN.—The stages in the breakdown of muscle glycogen (and of the glucose taken up by the muscle from the blood) are fully described on p. 846. At various stages in the "dissimilation" process, more high-energy phosphate bonds are generated which serve to restore creatine to creatine phosphate. Thus:

$$\text{Glycogen} \rightarrow \text{lower intermediaries} + {\sim}ph \text{ (as ATP)}$$
$${\sim}ph + \text{Creatine} \rightarrow \text{Creatine Phosphate.}$$

Under *anærobic* conditions, glycogen breaks down as far as pyruvic acid, which acts as an "intermediate hydrogen acceptor," *i.e.* it takes up hydrogen from reduced coenzyme-I (Co-I.2H) produced in an earlier reaction, and becomes lactic acid (cf. p. 849), *e.g.*

$$\underset{\text{(Pyruvic acid)}}{CH_3CO.COOH} + \text{Co-I.2H} \rightarrow \underset{\text{(Lactic acid)}}{CH_3CH(OH)COOH} + \text{Co-I}$$

It is clear, therefore, that skeletal muscle can function normally, *i.e.* contract and relax, for some time in the complete absence of oxygen. Fatigue, however, sets in rapidly.

Lactic Acid.—(i) If isolated muscle is stimulated to fatigue in an atmosphere of nitrogen (anærobically) lactic acid may accumulate up to concentrations of about 0·4% and equivalent amounts of glycogen disappear.

(ii) It must also be emphasized that in the intact person when the muscles are contracting *maximally* under conditions of natural blood flow, the oxygen supplies (though enormous) cannot be increased to the full extent demanded by the work of the muscle; in these conditions (in one sense ærobic, but in relation to need, really anærobic) lactic acid likewise accumulates in the

[1] Consult pp. 839–855. But "I pass slightly; I am not fond of repeating words like a parrot." (Gibbon.)

FATIGUE.—The cause of fatigue in directly stimulated isolated muscle is unknown. Fatigue resulting from stimulation of the nerve of a nerve-muscle preparation is due to some change at the motor nerve endings, possibly failure of the mechanism normally concerned with the release of the transmitter, acetylcholine (p. 513). The fatigue of muscular exercise in man is considered on p. 438.

Heat Production in Muscle.[1]—If a muscle is stimulated to contract *isometrically*, *i.e.* by pulling against a spring which it cannot overcome, the length of the fibres remains unaltered, but the tension in the muscle increases. All the energy of the chemical changes is converted into heat, which can be measured by means of a sensitive thermopile.

If the muscle is stimulated isometrically in the absence of oxygen (*anærobically*), heat is evolved. Some is evolved during the twitch, as shown in Fig. 258A; this constitutes the *initial heat*. A further much smaller amount is liberated for 1–2 minutes after the bout of activity is over, the so-called *anærobic delayed heat*. If the muscle is stimulated in the presence of oxygen (*ærobically*) the initial heat (during the twitch) is identical in amount with that liberated under anærobic conditions; but now a large amount of heat is also liberated after the end of contraction-relaxation, the *oxidative delayed heat*. These facts are expressed roughly quantitatively in the following Table (heat values are in arbitrary units):

	Initial Heat.	Delayed Heat.
Oxygen absent . .	3	0–1
Oxygen present . .	3	3

The course of heat production during isometric contraction followed by relaxation is well seen in Fig. 258A. The heat is derived from the energy liberated by the chemical changes detailed on pp. 428 *et seq.* (The effects of this liberated heat on *body temperature* are dealt with on p. 474.)

Work of Muscle.—When a muscle contracts *isotonically*, *i.e.* when it is permitted to shorten and move a weight, work is done. The possible mechanism responsible for the development of tension or of shortening was considered on p. 427.

The ratio $\dfrac{\text{energy converted into work}}{\text{total energy liberated}}$ represents the *mechanical efficiency*. In the case of isolated muscle under optimum conditions about 40% of the total energy of the chemical changes which occur is converted into work.

In the human subject lower figures are obtained: in an athlete the mechanical efficiency is about 25–30%; in an untrained person, it is 20–30%. With each kind of work there is an optimum *rate* of contraction at which the movement is carried out with a maximum efficiency (cf. Fig. 265 and p. 442).

[1] Hill, *Proc. roy. Soc. B.*, 1949, *136*, 195–254.

Circulatory Changes in Muscular Exercise.[1]—1. **Local Changes in Active Muscle.**—The circulatory adjustments in human muscles resulting from *local* activity can be determined by studying the volume and blood flow changes in the forearm (which is 85% muscle) during and after clenching the fist for varying periods and for a varying number of times.

(1) Blood Content Changes.—As the volume of muscle fibres is not decreased during contraction, any change in the volume of a muscle (recorded with the plethysmograph) must be due to alterations in its *content of blood* (or interstitial fluid). Muscle volume changes thus reflect changes (active or passive) in the calibre of the muscle vessels; they do not measure variations

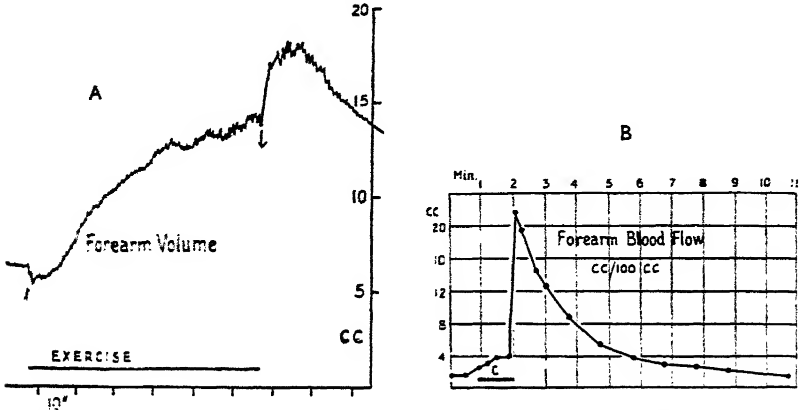


Fig. 260.—Effect of Clenching Fist on Circulation in Forearm. (Grant, *Clin. Sci.*, 1938, *3.*)

A. Forearm *volume* changes measured by simple plethysmography. Strong fist-clench maintained for 60 sec. (signal line, exercise). Time in 10 sec.
B. Forearm *blood flow* (in c.c./100 c.c. of forearm volume) measured by "occlusion plethysmography." C = strong clench of fist maintained for 1½ min. Time in minutes. Slight increase in blood flow during contraction; marked and sustained increase in blood flow after the contraction is over.

in the blood flow through the muscle which must be independently determined. Fig. 260, A shows that when the muscles contract, the *volume* of the forearm initially decreases; this is due to compression of the intramuscular veins and expulsion of their blood centripetally. If the contraction is maintained the volume of the forearm progressively increases; this is due to dilatation of arterioles and capillaries taking place owing to the action of locally formed products of metabolism (metabolites) and in spite of the handicap of raised muscular tension. When the muscles are relaxed their volume rises further, displaying the full vasodilatation which is now unimpeded by external compression. If the contractions are repeated several times (Fig. 261) it is found that each contraction temporarily diminishes the muscle volume (*i.e.* blood vessel calibre) to some extent; the volume mounts progressively, however, during each interval of relaxation, demonstrating that local vasodilatation is becoming progressively greater.

(2) Blood Flow Changes.—The resting *blood flow* in human muscle is very low, *i.e.* 1–4 c.c. per 100 c.c. of muscle. During a brief period of

[1] Bainbridge (ed. Bock and Dill), *Physiology of Muscular Exercise*, London, 1931. Schneider, *Physiology of Muscular Exercise*, London, 1933.

sustained contraction it may rise comparatively little; but during the subsequent relaxation it increases markedly (Fig. 260, B). These experiments show clearly that broadly speaking (as in heart muscle, p. 238) the skeletal muscle veins are emptied during contraction and that the arteries and veins fill during relaxation. The arterioles are dilated to a degree which is proportional to the severity and duration of the activity, the full measure of dilatation being manifest immediately activity comes to an end. Fig. 262 shows, for example, that after exercise of 4 minutes' duration the blood flow may be 33 c.c. per 100 c.c. of arm per minute, equivalent in this example to 30 times the resting value. The vasodilatation and high blood flow subside comparatively slowly, and 10–15 minutes may elapse before resting conditions are restored. This implies that the metabolites released during activity are relatively stable and are removed (or locally disposed of) very gradually. The vascular changes described are not significantly modified by depriving the limb of its sympathetic nerve supply; nervous factors obviously play little part in the local circulatory adjustments so far discussed. It is clear that it is easier to adjust local blood flow to local needs during *rhythmic* than in *static* muscular activity.

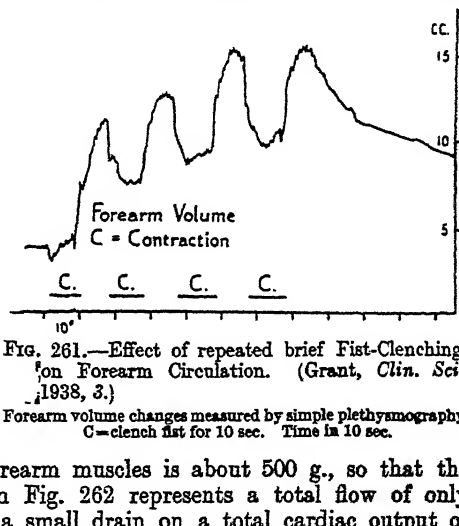


FIG. 261.—Effect of repeated brief Fist-Clenchings on Forearm Circulation. (Grant, *Clin. Sci.*, 1938, *3*.)

Forearm volume changes measured by simple plethysmography. C = clench fist for 10 sec. Time in 10 sec.

Such *local* adjustments can deal adequately with the needs of small masses of contracting muscle; the weight of the forearm muscles is about 500 g., so that the highest blood flow recorded in Fig. 262 represents a total flow of only 150 c.c. per minute, which is a small drain on a total cardiac output of 5000 c.c. When the muscular activity is very *widespread* and of maximal violence the local vasodilatation may be greater and anyway involves a far greater total blood flow. To take an extreme case: a blood flow of 30 c.c. per 100 g. to the entire muscle mass of 40 kg. (in man) represents 12,000 c.c. of blood flow per minute. To achieve muscle blood flows of this order of magnitude, widespread circulatory adjustments are needed.

2. General Circulatory Reactions.—The principal circulatory reactions may be thus summarized:

(1) VENOUS RETURN.—There is a greatly increased venous return which is due to: (i) vigorous muscular contractions, aided by the valves in the veins; (ii) increased depth and frequency of respiration; (iii) rise of capillary and venous pressure which is in part due to arteriolar dilatation.

(2) INCREASED CARDIAC OUTPUT (cf. p. 279).—The increased venous return is disposed of by a greatly increased cardiac output (on the right side) which is brought about as follows:

(i) By an increase in the *rate* of the heart to 150 or over per minute owing to:

(a) Impulses from the higher centres.

(b) Venous reflex, set up as a result of the raised pressure in the great veins (p. 272); this is probably the principal factor concerned and increases the heart rate in proportion to the rise in venous return.

(c) Rise of CO_2 tension acting directly (and reflexly) on the cardiac centre (p. 274 and Fig. 181).

(d) Rise of temperature acting on the cardiac centre and the sino-auricular node (p. 274).

(ii) By an increase in the output per beat (the stroke volume) up to perhaps 200 c.c.; the greater venous filling produces a greater stretch of the heart muscle fibres at the beginning of systole (p. 275).

The cardiac output may reach 30 or even 40 litres per minute, *i.e.* it is increased sixfold or more.

(3) The Left Side of the Heart receives the whole of the right cardiac output, and its output is equally increased.

(4) Redistribution of the Blood.—The blood is redistributed so that a greater proportion is sent to the active regions. The arterioles of the splanchnic area and skin are initially contracted, and the blood is diverted mainly to the muscles and heart. (When the body temperature rises in violent exercise the skin vessels are dilated again to facilitate heat loss (cf. p. 474).) The redistribution is effected as follows:

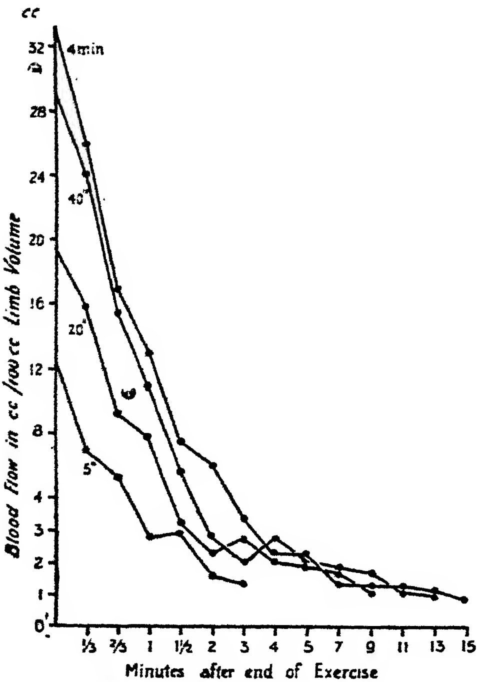


Fig. 262.—Arm Blood Flow following Exercise. (Drawn from data of Grant, *Clin. Sci.*, 1938, 3, 168.)

Blood flow measured by "occlusion plethysmography." Limb volume, 700 c.c. Initial resting blood flow, 0·9–1·4 c.c./100 c.c. Exercise consisted of strong fist-clench for 0·5 sec., followed by interval of 1 sec. Duration of period of exercise (shown on each curve): 5", 20", 40", 4 min.

(i) The *vasomotor centre* is stimulated by (a) impulses from the higher centres; (b) perhaps reflexly from the active muscles; (c) by the increased CO_2 tension (p. 309). It will be recalled that the vasomotor fibres pass chiefly to the splanchnic and cutaneous arterioles.

(ii) *Peripherally*, (a) *metabolites* are the main factors involved and they dilate the arterioles in the muscles (p. 433); (b) *adrenaline*, which is secreted in times of stress (p. 731), constricts the arterioles which have a marked vasomotor supply (skin and splanchnic area) and directly dilates muscle vessels.

The blood flow to the *brain* is perhaps increased as it is chiefly dependent on the general level of blood pressure, which is raised. The blood supply

to the *heart* is enormously increased for reasons discussed in detail on p. 238.

The entire circulatory system is coordinated to one main end, *i.e.* to increase as far as possible the blood flow through the active tissues. This result is attained by the nice cooperation of a large number of factors: the respiratory and muscular "venous pumps," the heart, the vasomotor nerves, and peripherally acting chemical dilator agents.

(5) Blood Pressure.—The *blood pressure* rises, *e.g.* to 160 or 180 mm. Hg. As the output of the heart increases threefold or even sixfold, one would expect the blood pressure to rise in the same proportion. The fact that the observed rise of blood pressure is relatively small shows that the total *peripheral resistance in exercise is decreased, i.e.* the dilatation of the muscle vessels far outweighs the constriction of the splanchnic area and skin.

(6) Capillaries.—The muscle capillaries are dilated, and many which were previously closed become patent owing to the action of metabolites (p. 320). This does not alter the volume of blood flowing through the muscle in unit time, but it decreases the linear velocity and, as has been pointed out (p. 413), this provides more time for the all-important gaseous interchanges between the blood and the tissues.

Hæmo-Respiratory Changes.—The principal points to be considered are: (i) pulmonary ventilation; (ii) oxygen consumption; (iii) CO_2 output; (iv) ratio of CO_2 output/O_2 consumption, *i.e.* the respiratory quotient; (v) changes in the blood, especially its reaction and gas-carrying properties. The changes must be reviewed (i) in *moderate* and in *very severe* exercise, and (ii) during the exercise and after it has come to an end. We will further have to consider broadly how the vast amounts of oxygen which are needed by the active muscles are supplied to them and how the CO_2 formed is removed and finally expelled from the lungs.

The size of the respiratory problem is indicated by the following data: the O_2 consumption in severe exercise may be raised from 250 c.c. to 4000 c.c. per minute (*i.e.* sixteenfold); the CO_2 output in some phases of violent exercise may go up from 200 c.c. per minute to nearly 10,000 c.c. (*i.e.* about fiftyfold).

Varieties of Exercise.—(1) The term "very severe exercise" is used with reference to muscular activity which by reason of its severity can only be kept up for a very short time; examples of this are a 100-yards or quarter-mile race at top speed; at the end of such a race the runner is completely exhausted. From the physiological point of view the characteristic feature of this form of exercise is the inability of the hæmo-respiratory systems to supply the muscles *during* the period of exercise with all the oxygen they require for their tremendous level of activity. A so-called *oxygen debt* is incurred (p. 439), *i.e.* a large volume of oxygen has to be absorbed after the exercise is over to dispose of metabolites which accumulated in the muscles during activity because of their relatively anoxic state.

(2) "Moderate exercise" is of a kind that can be kept up for long periods, *e.g.* vigorous walking at 5 miles per hour or steady running. Such exercise may involve a considerable measure of exertion and necessitate extensive hæmo-respiratory adjustments. Its outstanding physiological feature is the ability of the body to supply the active muscles with practically all the oxygen that they require *immediately* it is wanted and to a degree which is

proportional to the level of activity. These points are further discussed below.

METABOLIC STUDIES ON EXERCISE IN MAN.[1]—The resting oxygen consumption and the CO_2 output are first determined by means of the Douglas bag technique. The expired air is then collected *during* and *after* exercise, and the gaseous exchanges again determined. The metabolism over and above the resting level can thus be calculated, and is termed the *excess metabolism of exercise.*

Moderate Exercise.—(1) PULMONARY VENTILATION.—The pulmonary ventilation is increased to a degree which is proportional to the severity of the exercise. Thus when walking at 2, 3, 4, and 5 miles per hour the pulmonary ventilation rose to 18·6, 24·8, 29·0, and 60·9 litres per minute respectively (p. 405). The increased breathing is brought about by appropriate stimulation of the respiratory centre:

(i) by impulses from the higher centres produced by the emotional tension before the exercise begins;
(ii) by the raised CO_2 tension of the blood;
(iii) by any rise of body temperature that may occur;
(iv) reflexly from the engorged great veins and from the active muscles themselves.

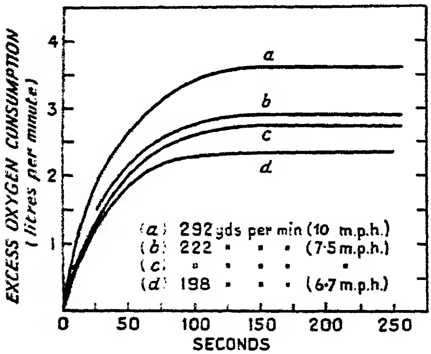


FIG. 263.—Relation of Oxygen Consumption to Severity of Exercise (Speed of Running) in Man. (A. V. Hill, *Muscular Movement.*)

Ordinate: excess oxygen consumption above resting level (litres per minute).
Abscissa: time in sec. after onset of running. More oxygen is consumed at higher speeds; the oxygen consumption rises gradually and reaches its maximum steady value after about 2 minutes.

Very little lactic acid accumulates in the blood, as it is disposed of in the muscles and other organs about as rapidly as it is being formed; arterial anoxia does not develop, as the blood flowing through the lungs is adequately oxygenated. The H^+ ion concentration of the blood only rises slightly from H_2CO_3 retention.

The most important single factor controlling respiration is the rise of CO_2 tension; the ventilation increases in direct proportion to the severity of the work performed. After the exercise is over the breathing rapidly settles down to the resting level as the excess CO_2 is eliminated.

(2) OXYGEN CONSUMPTION.—This is increased in proportion to the severity of the exercise. Thus when walking at 2, 3, 4, and 5 miles per hour the O_2 consumption rose to 780, 1065, 1595, and 2543 c.c. per minute. Fig. 263 shows that in running at 7, 7·5, and 10 miles per hour the O_2 intake rises gradually, reaches its peak after about 2 minutes, and then remains steadily at this high level; at 10 miles per hour the O_2 intake per minute was 3·6 litres per minute. It is convenient to consider here the means by which most of

[1] Hill, *Muscular Movement in Man*, London, 1927. Grodins, *Physiol. Rev.*, 1950, *30*, 220. Gemmill, *ibid.*, 1942, *22*, 32. Comroe, *ibid.*, 1944, *24*, 319.

this enormous amount of oxygen is supplied to the active muscles. The main reactions concerned are as follows :

(i) *Breathing.*—The increase in pulmonary ventilation already described introduces large amounts of fresh air and oxygen into the lungs (and drives out CO_2).

(ii) *O_2 Uptake in the Lungs.*—Large amounts of oxygen are taken up from the lungs by the blood.

(*a*) The mixed venous blood arrives at the lungs in a more reduced form. Normally at rest the venous oxygen content is 14 c.c.%, the amount taken up from the alveolar air is 5 c.c.% and the arterial oxygen content is 19 c.c.%. In hard work the venous blood may contain 7 or even as little as 3 c.c.% ; in becoming normally arterialized it takes up 12–16 c.c.% (cf. p. 413). The oxygen intake may by this means alone increase more than threefold.

(*b*) The cardiac output is also increased about fourfold or more (*e.g.* up to 30 litres per minute (pp. 279, 433)).

The oxygen uptake in the lungs may in these ways be increased to the requisite amount, *e.g.* sixteenfold (from 0·25 litre to 4 litres per minute).

(iii) *Supply of Oxygen to the Tissues.*—A great blood supply (and therefore oxygen supply) to the muscles results from the large left heart output and the redistribution of the blood in the systemic circulation (p. 434).

(iv) *Removal of Oxygen by the Tissues.*—This is effected as follows :

(*a*) Dilatation and increase in the number of the capillaries in the muscles (p. 320) slows the rate of the blood flow and allows more time for gaseous exchanges.

(*b*) Low oxygen tension in the tissues allows oxygen to diffuse more readily and to a greater extent out of the blood.

(*c*) High CO_2 tension and raised temperature increase the extent and rate of dissociation of oxyhæmoglobin (p. 411).

The blood which leaves the tissues is thus very extensively reduced, and the mixed venous blood may contain as little as 3 c.c.% oxygen instead of the usual 14 c.c.%. Experimentally, the O_2 consumption of active muscle may rise thirtyfold. This could be achieved by an increased blood flow (*e.g.* $\times 10$) and an increased utilization of oxygen (*e.g.* $\times 3$) ; it is known that the muscle blood flow in exercise may increase even up to thirtyfold, so that an ample margin is available.

It will be noticed again how every part of the body makes its contribution to the general effort. The pulmonary "bellows" supply the oxygen. The heart and vasomotor mechanism send the blood mainly to the parts which need it. The tissues abstract as much oxygen as they can from the blood which is supplied to them.

The O_2 consumption falls *rapidly* to its resting level when the exercise is ended.

(3) CO_2 Output.—Data for CO_2 output are given in the Table on p. 405. The ratio of CO_2 elimination to oxygen consumption can be judged from the data for the respiratory quotient given below. The elimination of the large amounts of CO_2 formed in the body is effected in a manner analogous to that described for oxygen.

(4) Respiratory Quotient.—Special attention has been paid to the R.Q. of the total metabolism or of the excess metabolism of exercise in the hope (possibly the vain hope, cf. p. 373) of obtaining information about the kind

of foodstuffs which are used by muscles during activity. The R.Q. proves to be about the same as the pre-exercise value (*e.g.* 0·85). Thus when walking at 2, 3, 4, and 5 miles per hour, the R.Q. varied from 0·85 to 0·9; during the previous resting period it was 0·8 to 0·88 (p. 405). If these results mean anything they suggest that the body uses the various foodstuffs in exercise in roughly the same proportions as at rest.

On a diet consisting mainly of *fat*, if the exercise is *prolonged*, the R.Q. falls to about 0·7, suggesting that fat is being used as the predominant fuel. The fat is *not* burnt directly by the muscles, but is first transformed in the liver into carbohydrate or ketone bodies which are then sent to the muscles to serve as the source of energy.

(5) Blood Reaction.—Lactic acid does not escape from the muscles into the blood. The H^+ ion concentration of the blood only rises slightly from accumulation of H_2CO_3, and there is consequently no marked respiratory distress.

(6) Cardiac Output.—In running at the rate of 9 miles per hour, an oxygen intake of about 4 litres per minute is attained and can be kept up for half an hour. This, as explained on p. 279, involves a cardiac output per minute of about 30 litres, which is the highest obtained in the human subject when breathing air. If oxygen is breathed during the period of exercise, the oxygen consumption may be as high as 5·9 litres per minute, and the cardiac output is then over 40 litres per minute. This latter figure probably represents the limits of bodily reaction.

The other circulatory changes are described on pp. 432 *et seq.*

(7) Fatigue.—In steady prolonged exercise, fatigue is due to a number of ill-understood factors; in the main it is attributed to changes in the brain resulting from slight anoxia and increased H^+ ion concentration. Afferent impulses set up in the active muscles (in part perhaps by the local physico-chemical changes) give rise to discomfort and contribute to the sense of weariness. Some of the stiffness may be due to swelling of the muscle from accumulation of exuded fluid from the blood (p. 19).

Severe Exercise.—(1) Pulmonary Ventilation.—The characteristic feature of this type of exertion, which is necessarily of relatively brief duration, is that the breathing remains much above the resting level for a prolonged period *after the exertion is over.* In the case of a man who ran 225 yards in 23·4 seconds, the pulmonary ventilation returned to normal in 27 minutes; after a quarter-mile race followed by severe gymnastics, in 44 minutes; after "standing-running"[1] for 4 minutes (breathing oxygen), in 87 minutes. In the case of a 100-yards sprint, the subject may scarcely draw breath during the race, but marked dyspnœa develops later.

(2) Lactic Acid Formation.—During very violent exercise, owing to relative anoxia of the muscles, lactic acid accumulates in the muscles and diffuses out into the blood stream and throughout the body fluids.

The resting level of blood lactate is 10–20 mg–%. As a result of violent exercise the level may rise to 100 or even to 200 mg–%. As lactate is freely diffusible, the blood lactate level represents the concentration in the muscles, the interstitial fluid, and possibly in the intracellular fluid generally. If the total volume of body fluid (*i.e.* 50 L) contains a lactate concentration of

[1] *i.e.* standing in one place and carrying out rapid movements of the legs as in running.

200 mg-% (*i.e.* equal to that in the blood), the total lactate accumulation is 100 g. After the exercise is over this lactate is disposed of; the blood lactate concentration steadily diminishes and reaches the normal level after a variable period of time, sometimes after as long as 60 minutes.

(3) Respiratory Quotient.—The respiratory quotient of excess metabolism during the period of exertion first rises above 1 and may reach 1·5 or 2; the maximum figure is usually attained shortly after the end of exercise. During the recovery period the R.Q. falls below normal, *e.g.* to 0·5.

These changes can be readily accounted for. The lactic acid which, as we have just noted, accumulates in the plasma during violent exercise is buffered as usual by the bicarbonate.

$$HL + NaHCO_3 \rightarrow NaL + H_2CO_3 \rightarrow CO_2$$

This reaction results in the liberation of large amounts of CO_2, *without any equivalent utilization of oxygen.* In addition, various foodstuffs are burnt in the muscles, oxygen being used and CO_2 evolved with an R.Q. of 1 or less. When the R.Q. of excess metabolism is 2, for each 1 molecule of CO_2 resulting from oxidation processes in the muscles, at least another 1 molecule of CO_2 is evolved from the $NaHCO_3$ of the plasma. The extent to which the R.Q. exceeds 1 is a good index of the intensity of the exertion. The low R.Q. following the exercise is due to CO_2 being *retained in the blood to re-form the bicarbonate* (p. 440).

(4) **Recovery after Severe Exercise.**—After severe exercise the oxygen consumption (like the pulmonary ventilation) remains initially far above the resting level; thus following violent standing-running (Fig. 264) the O_2 intake *per minute* was at the rate of 1800 c.c. at 30 seconds, 1250 c.c. at 50 seconds, 750 c.c. at 100 seconds, and 500 c.c. at 140 seconds; (the resting level was 250 c.c. per minute). The O_2 intake declines slowly further but may not return to resting level for 30–120 minutes.

The volume of O_2 used after the exercise is over in excess of the resting O_2 consumption for the same length of time is called by Hill the *oxygen debt*, or, better perhaps, the *recovery oxygen.* It can be measured as follows:

(i) the resting O_2 consumption is determined.

(ii) The post-exercise oxygen consumption is measured until it has fallen to its pre-exercise value; as a rule this occurs in 30 minutes. The resting oxygen consumption in (say) 30 minutes is then deducted from the post-exercise O_2 consumption for the same time. The difference is the O_2 debt. Oxygen debt figures as high as 15–18 litres have been observed. These results can be readily accounted for. During very violent exercise all the tremendous circulatory and respiratory reactions prove inadequate to supply the active muscles with their full O_2 requirements. As has been repeatedly emphasized, the active muscles, in spite of their large O_2 uptake are contracting in a sense anærobically. Much lactate undoubtedly accumulates in the body (up to 100 g., *supra*) and probably other products of muscular metabolism as well. The recovery oxygen is used to dispose of the various waste products which have accumulated during the bout of violent activity.

(iii) During the first few minutes of the recovery process there is no decrease in the amount of lactate in the plasma nor, presumably, in the muscles. It seems, therefore, that metabolites other than lactate are disposed of first. The oxygen used for this purpose is termed the "alactic acid debt."

(iv) When this has been dealt with the lactate is got rid of gradually in the following way:

(a) The lactate of the muscles is first disposed of; *i.e.* it is reoxidised to pyruvate and then dissimilated to CO_2 and water or perhaps partially reconverted into glycogen (p. 849).

(b) The sodium lactate of the blood is in the ionized form as Na^+ and lactate′; some of this lactate is now taken up by the liver and is there rebuilt into glycogen.

(c) The rest diffuses into the muscles (as the local lactate concentration

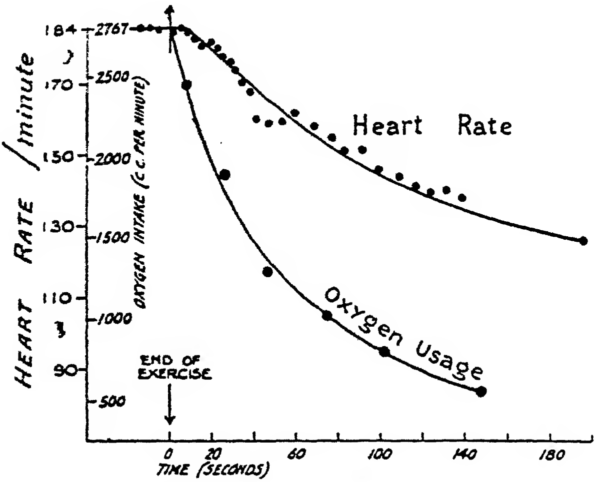


Fig. 264.—Recovery Phase after Severe Muscular Exercise (Standing-Running).

Abscissa = time in seconds. Ordinate: heart rate per minute; O_2 intake in c.c. per minute. The resting values for heart rate and oxygen consumption were 70 beats and 250 c.c. per minute.]
End of exercise shown by arrow.
Note that neither the O_2 consumption nor the heart rate had returned to normal 140 sec. after the end of exercise. *Total* excess O_2 consumed after the exercise is over is the *oxygen debt*. (After Lythgoe and Pereira.)

falls) and other organs to be oxidized and then dissimilated or converted into glycogen. Such lactate as is oxidized in the tissues yields CO_2; (it is perhaps worth emphasizing that lactate cannot be oxidized in the blood stream.

(d) The CO_2 thus formed, on entering the blood is *retained* there to a considerable extent instead of being blown off in the lungs. It unites with the Na^+ of the plasma which was previously combined with lactate (as sodium lactate) to *re-form* the bicarbonate which has been depleted during the period of exercise.

(e) The CO_2 output from the lungs is thus considerably less than the oxygen consumed, and in this way the low R.Q. which follows violent exercise is accounted for.

(5) Blood Reaction.—Owing to the flow of lactic acid into the blood there is an increase in the H_2CO_3 and a fall in the bicarbonate content of

the plasma. These two factors produce a demonstrable increase in the H^+ ion concentration. This *acidæmia* stimulates respiration and increases the pulmonary ventilation to an extent which is sufficient to give rise to subjective symptoms of distress or *dyspnœa*. The acidæmia is still present, of course, at the end of this severe type of exercise. The ventilation,

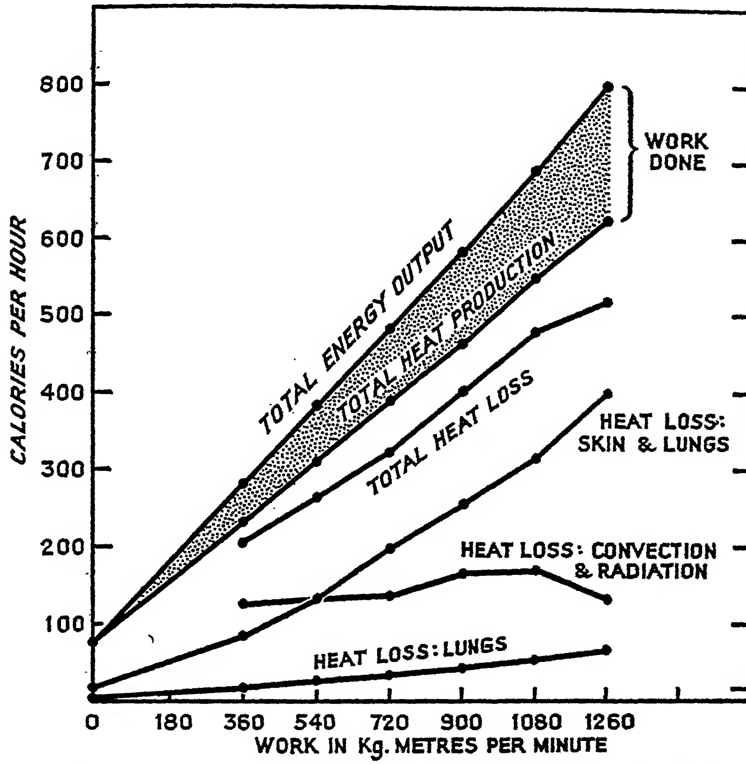


FIG. 265.—Analysis of Energy Exchanges in Muscular Exercise. (After Nielson, *Skand. Arch., Physiol.*, 1938, *79*, 193.) (Cf. Fig. 278.)

Ordinate: Calories per hour.
Abscissa: Work in kg. metres/minute.
The work in each case lasted 60 minutes and was of increasing degree of severity.

therefore, does not return to normal for a prolonged period, *e.g.* for half an hour or more after the exercise is over; as already pointed out throughout this time the oxygen consumption (and the cardiac output and blood flow to the muscles) remain above the resting level.

After the exercise is over the normal blood reaction is restored (i) by blowing off the excess CO_2; (ii) by restoring the bicarbonate of the plasma to its original level, as described above.

Very little lactate is excreted in the urine, and the amount thus eliminated is no index of the amount of lactic acid formed.

It must be emphasized that strenuous exercise does *not* depend on *concurrent* oxidation; lactic acid and other waste products are formed in amounts far greater than the maximum oxygen intake can cope with. The chemical processes (partly ærobic, but to a varying extent also anærobic) going on in the muscles yield energy which enables activity to be continued. The excess lactic acid in man is temporarily buffered in the tissues and in the blood stream, accumulates throughout the body fluids, and is disposed of *after* the exercise is over. We are thus permitted to do work far in excess of the greatest oxygen supplies which the heart and lungs can provide at the time, and we accumulate lactic acid and other substances. We metaphorically overdraw our oxygen account during the exercise and repay the debt later. The size of the repayment during the recovery period is, of course, equal to the oxygen debt incurred during activity. The acidæmia which is still present at the end of exercise serves so to say as a guarantee that the ventilation will be adequate to supply all the oxygen necessary for the recovery process.

SECOND WIND.—This phenomenon is familiar to athletes; during a long race initial dyspnœa supervenes, but after a while this passes away and the breathing becomes more comfortable. It has not yet been satisfactorily explained.

Heat Exchanges.—Fig. 265 shows the changes occurring during exercise of increasing degrees of severity. Total energy output increased in direct proportion to the increase in the severity of the work. With the hardest type of work studied (1260 kg./min.) the total energy output rose to 800 Cal. per hour as against a resting value of less than 100 Cal. per hour. Only about one-quarter of this energy is converted into mechanical work, and about three-quarters (600 Cal. in the example quoted) was given off as heat. The change in body temperature depends on the relationship between the rate at which heat is formed and the rate at which heat is given off from the body. With hard work heat production substantially exceeds heat loss with the result that the body temperature rises. The lower three curves in Fig. 265 show the ways in which the extra heat is given off. Heat loss by convection and radiation alters negligibly, and there is a small increase in heat loss from the lungs; in the main, heat loss is effected by *increased secretion and evaporation of sweat.*

ANOXIA (OXYGEN LACK) [1]

The term *anoxia* is employed to describe oxygen lack in the body from any cause. Anoxia occurs commonly and complicates many diseases as a secondary factor. It is important to obtain a clear idea of the effects of anoxia and the various conditions which may produce it.

Effects of Anoxia.—1. NERVOUS SYSTEM.—When anoxia is very rapid and severe, *loss of consciousness* may occur without any warning, as when the oxygen supply is cut off from an airman at a great height or when a miner puts his head into a cavity full of methane gas and drops as though

[1] See Barcroft, *Respiratory Function of the Blood*, Cambridge, 1914; 2nd edn. pt. i. *Lessons from High Altitudes*, 1925; *Lancet*, 1920, ii, 485. Haldane and Priestley, *Respiration*, Oxford, new edn., 1935. Van Liere, *Anoxia*, Chicago, 1942. Armstrong, *Aviation Medicine*, 2nd edn., London, 1943. Monge, *Physiol. Rev.*, 1943, *23*, 166.

felled by a blow. When the miner recovers on breathing pure fresh air for a few seconds, he really believes that he has been knocked down, and acts accordingly, assaulting anyone in his vicinity. If anoxia develops more gradually, the intellect and the senses become dulled without the person being subjectively aware of what has happened. When the anoxia is relieved, the sudden increase in the power or in visual acuity comes as a great surprise. Not infrequently, however, the symptoms are temporarily aggravated on breathing normal air.

The symptoms of severe anoxia may resemble those resulting from an overdose of alcohol: headache, depression, apathy and drowsiness, or excitement and general loss of self-control. The subject will sing and shout, or burst into tears for no apparent reason. He may be extremely quarrelsome or insolent. If in any danger, he may refuse to take such measures as are obviously essential for his safety. He feels quite confident that his mind is clear and that his judgment is sound, in spite of his dangerous behaviour. Memory is impaired, and appreciation of time is altered. He may think he has gone quite a long way when he has only stumbled a few yards.

The effects of acute severe anoxia are well brought out in the following record.[1] At 2.53 a.m. one November morning an American aviator (L.) took off into the darkness for the usual weather flight ("Dawn Patrol"). It was customary to climb to 16,000 feet, levelling off every 1500 feet for a short period. Owing to some instrumental fault he really climbed 2 feet for every one read on the altimeter. The flight proceeded normally to an altitude of "8000" feet (actually 16,000), when L. began to feel weak, dizzy, and slightly confused; there was respiratory distress and a sense of oppression in the chest; his flying co-ordination was also imperfect. He suspected he might be in need of oxygen, but as he had never used oxygen at lower altitudes than 12,000 feet and his altitude was only "8000" feet, he thought that could not be the trouble. He was maintaining wireless contact with the ground, but at this stage could not find the proper words to use. All this did not alarm him in the least. Though weak and bewildered he began to ascend further; in the meantime he dropped his microphone. He fumbled all over the cockpit in the dark searching for it, but did not think of turning on the light. After five minutes' frantic search he gave up the hunt; the microphone was subsequently found lying on his lap. By this time he had reached "10,000" (actually 20,000) feet; the plane now went into a spin. Not the least disturbed, he climbed again, fell once more into a spin, and repeated the performance several times. None of these experiences alarmed him but merely increased his determination to reach his objective of "16,000" feet. Breathing was now difficult and accompanied by a sense of suffocation; muscular co-ordination was defective and there was considerable emotional upset. He was angry and possessed of an acute sense of failure: other pilots could reach 16,000 feet without difficulty and here he was unable to attain "12,500" feet without a spin. Grim determination drove him to spend over an hour climbing and spinning; from anger and shame tears were streaming down his face almost all the time. All sense of fear or of passage of time was gone; vision became limited to the nose of the plane and the instrument panel; flying became semi-automatic. "He was fighting mad till he became so tired that he

[1] Leedham, *J. Aviation Med.*, 1938, *9*, 150.

began to pass out." After being an hour in this twilight state he woke up in a full throttle dive towards the ground. The improved oxygen supply brought him once more to his senses at "4000" feet and he began to climb again, though he only had enough petrol in the emergency tank for 15 minutes' use. At about this time (at the lower altitude) he became more fully conscious, and a realization of what had been happening dawned on him. After alighting, he noted fatigue, poor appetite, tremor, and loss of confidence; but he was quite normal again after 24 hours.

In anoxic states understanding is impaired more than sensation; the subject sees without knowing what he is looking at. He finds it difficult to understand printed or written words. Pain is dulled: a miner poisoned with CO gas readily burns himself with his lamp. Finally, each sense is lost suddenly, hearing being retained longest. There is great muscular weakness and easy fatiguability. Progress is made with extreme slowness. The mountaineer "may think twice before turning over in bed." Sudden paralyses of lower or upper limbs may occur.

Very trifling changes in the environment may make a great difference to the affected person. The importance of attending to the minutiæ of nursing comfort in the case of anoxic patients is obvious.

Acute anoxia occurs clinically in some cases following pneumonectomy. The patient becomes disorientated and does not know where he is; he climbs out of bed and may become violent or quite unmanageable.

2. Digestion.—There is nausea, loss of appetite, and vomiting. At 26,700 feet, the members of the Everest expedition could with difficulty bring themselves to eat meat at all. Chocolates and biscuits were consumed as a duty, and only soup and coffee were taken with relish. (On the mountain-side, great thirst was a striking feature owing to dryness of the mouth and throat resulting from the marked loss of fluid from the respiratory passages in the dry cold air.)

3. Circulation.—There is at first an increase in the frequency of the heart and probably in the minute output. A small rise of blood pressure may occur.[1] These results are due to anoxia acting (mainly reflexly) on the vasomotor and cardiac centres in the medulla. The blood flow to the splanchnic area and skin is reduced and more blood is diverted to the brain and the heart muscle. Later, the force of the heart dies down, though the frequency may increase further. In the Everest expedition the pulse rate during motion was 160–180 per minute, or even more, of good volume and regular. All the men who went above 27,000 feet developed dilated hearts, which returned to normal at lower altitudes in 1 to 3 weeks. Owing to the imperfect oxygenation of the blood, *cyanosis* develops (for full discussion see p. 452). The mucous membranes may appear black or leaden grey according to the amount of blood in the capillaries.

4. Respiration.—The effects on the breathing have already been fully considered (p. 398). The breathing is increased in rate, but tends at first to be shallow and periodic in character.

5. **Delayed Effects.**—Beside the immediate results of anoxia, numerous delayed effects appear, depending on the length and severity of the exposure. An inspector of mines who had been in an anoxic atmosphere, on returning to the surface, first shook hands cordially with all the bystanders. The

[1] Cf. p. 273, p. 310, and Fig. 243 for effects of acute anoxia on the circulation.

doctor offered him his arm; he regarded this as an insult, and there and then took off his coat and challenged him to fight. Mountain sickness with typical nausea, vomiting and depression, may ensue 8–12 hours after the exposure, perhaps even after the descent. A slight degree of oxygen lack may therefore act insidiously and only reveal its effects after a long latent period. This conclusion is of importance when we consider that mild anoxia is fairly common clinically. Following prolonged exposure to severe oxygen lack, very formidable after-effects occur. Anoxia "not only stops the machine but wrecks the machinery" (Haldane).

In CO poisoning, consciousness may not be regained for many hours after the condition of the blood has been restored to normal; or coma may persist and pass into death. Convulsions may occur at intervals owing to injury to nervous elements, and paralyses of various sorts may follow. The mental condition may be temporarily impaired or the subject may pass into a state of dementia. Pneumonia may set in after a day or two; this complication is an index of the generally precarious condition of the whole body.

6. **Acclimatization.**—Given a slow onset of anoxia and favourable circumstances, the various compensatory activities of the body are brought into play. At moderate heights, the early unpleasant symptoms may pass away completely. At Pike's Peak (14,100 feet), the general discomfort, the fatigue, and the periodic breathing passed off in 10 days, and all the subjects felt and looked extremely well. Similarly, in the anoxia of disease we may expect to find the symptoms masked for a long time, only to be strikingly displayed when compensation breaks down.

Acclimatization is effected as follows:

(i) The *kidney* excretes urine of an alkaline reaction and with a low NH_4^- content; the alkalæmia is thus effectively compensated.

(ii) The *pulmonary ventilation* is consequently permitted to proceed at a sufficient level to maintain a much higher alveolar oxygen tension than otherwise would have been the case.

(iii) The red bone *marrow* is stimulated by the chronic anoxia; red cells are formed and turned out into the circulation in increased numbers; the reticulocyte count (*i.e.* the proportion of young red cells) rises. The red cell count and the hæmoglobin concentration rise (the latter may go up to 115–150% on the Haldane scale (p. 175)); the *oxygen-carrying capacity* of each c.c. of blood is thus increased.

The changes during acclimatization in pulmonary ventilation, alveolar CO_2 tension, and hæmoglobin concentration are shown in the Table below.

Altitude.	Days after Arrival.	Alveolar CO_2 mm. Hg.	Relative Pulmonary Ventilation.	Hb%. (Haldane) Scale
Sea level .	—	40	100	100
6000 feet .	1	37	108	100
	3	35	115	101
14,000 feet .	1	32	125	103
	5	29	138	109
	20	27	148	116

Problems of High Altitude Flying.[1]—At great heights oxygen must be provided or pressurised cabins used.

(1) Oxygen is supplied from cylinders of liquid oxygen equipped with means for converting the oxygen into the gaseous form. A face-piece is fixed over the mouth and nose, through which a current of oxygen flows constantly; about half the oxygen delivered in this way is wasted because it reaches the face-piece during expiration, but the method has the advantage of simplicity of apparatus. Donati, for example, attained by this means a height of 47,358 feet. At 45,000 feet the barometric pressure is 110 mm. Hg (Fig. 244); this must also represent the total pressure of the alveolar air. Of this pressure in the alveoli water vapour represents 47 mm. Hg and CO_2, say 40 mm. The residual pressure, 23 mm. Hg, is all that is available for alveolar oxygen pressure, *even when pure oxygen is breathed.* Actually the anoxia stimulates respiration and decreases the CO_2 pressure (the inevitable result of hyperpnœa at rest); this allows a corresponding increase in the oxygen pressure, possibly to about 40 mm. Hg. At such a pressure consciousness would be retained by a person in good training. If the barometric pressure was 130 mm. there would be little danger so long as pure oxygen was breathed, but 110 mm. seems to be the lowest tolerable barometric pressure, using the type of apparatus described above.

(2) To ascend to greater heights and for civilian flying where oxygen inhalation is inconvenient, an air-tight cabin is used which contains air of normal composition, kept at normal barometric pressure irrespective of the height attained.[2]

Types of Anoxia.[3]—Four types of anoxia may be recognized:

(1) *Anoxic type*: the *tension* of oxygen in the arterial blood is lower than normal, and consequently the hæmoglobin is not saturated with oxygen to the normal extent.

(2) *Anæmic type*: the oxygen tension in the arterial blood is normal, but the quantity of *functioning hæmoglobin* is too small.

(3) *Stagnant type*: the arterial blood is normal in oxygen tension and oxygen content, but it is supplied to the tissues in *insufficient amounts* because of a decrease in the blood flow.

(4) *Histotoxic type*: the *tissue cells are poisoned* and cannot make effective use of the oxygen supplied to them.

We may now discuss each form more fully.

1. Anoxic Type.—This very serious form of anoxia occurs characteristically at high altitudes. The tension of oxygen in the alveolar air is too low; the tension and content of oxygen in the arterial blood are consequently lowered too.

When the oxygen tension in the blood falls, breathing is stimulated and consequently (as already explained) excessive amounts of CO_2 are washed out from the blood. As a result oxyhæmoglobin tends to retain its oxygen (Fig. 249), and so the tissues are starved of oxygen independently of

[1] Fulton, *Aviation Medicine in its Preventive Aspects*, London, 1948. Ivy, *Fed. Proc.*, 1946, *5*, 319. For effects of acceleration in relation to aviation see Wood *et al.*, *Fed. Proc.*, 1946, *5*, 327.

[2] For dangers of decompression sickness at high altitudes see p. 426.

[3] Argyll Campbell, and Poulton, *Oxygen and Carbon Dioxide Therapy*, London, 1934. Barach, *Inhalational Therapy*, Phila., 1944.

the amount present in the blood. The *velocity* of the oxidative processes seems to depend to some extent on the *pressure* head at which oxygen is supplied.

The tissues, therefore, are hampered in three ways:

(*a*) The tension of oxygen in the blood is low, and so the rate of tissue oxidation is diminished.

(*b*) There is less oxygen present in the blood, so that any activity demanding a large oxygen consumption is impossible.

(*c*) The oxygen which is present in the blood is not readily available to the tissues, owing to the low CO_2 tension hampering the dissociation of oxyhæmoglobin.

Barcroft illustrated this very well in his experiment in a respiration chamber where he was exposed to low oxygen tensions. Normally, his oxygen capacity was 17·8 c.c.%, the hæmoglobin 96% saturated, and the oxygen content of the arterial blood was 16·9 c.c.%. At low oxygen tensions his hæmoglobin became only 84% saturated, but as his hæmoglobin concentration rose, his O_2 capacity went up to 20·1 c.c.%; his arterial O_2 content was therefore still 16·9 c.c.%. Although the volume of oxygen present in his arterial blood was the same as normal, severe symptoms were present in the respiration chamber (headache, vomiting), because the oxygen *tension* was subnormal.

Clinical Causes of Anoxic Anoxia.—(1) ALTERATION OF THE ALVEOLAR EPITHELIUM.—If the walls of the alveoli are swollen and œdematous, or covered by a layer of fibrinous exudate, diffusion of oxygen into the blood cannot readily occur. The arterial blood does not attain the same oxygen tension as the alveolar air and is therefore undersaturated. This is well seen as a result of poisoning with irritant war gases. The alveoli become sodden with fluid, and cyanosis is present in spite of deep breathing. Administration of oxygen improves the colour and the general condition of the patient, because at the higher alveolar oxygen tension the gas diffuses more rapidly into the blood. In heart failure with *œdema of the lungs*, anoxia is caused in the same way.

(2) PARTIALLY UNVENTILATED AREAS OF LUNG.—(i) SHALLOW BREATHING.—Keith suggests that the various regions of the lungs do not open out equally and simultaneously during inspiration, but part by part, "like a lady's fan." The alveoli nearest the moving parts of the chest wall and the diaphragm expand first, and to the greatest extent; the other regions follow suit later, and to a less extent. In shallow rapid breathing certain parts of the lungs are well opened up, and because of the frequency of the breathing receive more than their normal amount of fresh air. The alveolar CO_2% in these regions falls, and consequently a greater amount of CO_2 is washed out from the blood. Similarly, the oxygen in these alveoli may rise to 16 or 18%, but this is of practically no value to the blood. As arterial blood is 95% saturated at the usual alveolar oxygen tension, it takes up a negligible additional amount of oxygen if the pressure is raised. Many parts of the lung, however, are considerably underventilated, as the breathing is not deep enough or prolonged enough for them to be opened up. Less CO_2 is given off in these areas; because of the low oxygen tension in these alveoli, the blood leaves them considerably undersaturated with oxygen. If the *total* alveolar ventilation is normal, the CO_2 is eliminated in adequate amounts; but not enough

additional oxygen is taken up in the overventilated alveoli to compensate for the incomplete saturation which occurs in the underventilated alveoli. As a result, the blood which leaves the lungs is deficient in oxygen.

In bronchopneumonia and asthma small patches of collapse occur and the blood leaving them is underarterialized.

(ii) LOBAR PNEUMONIA.[1]—In this disease a number of factors contribute to the development of anoxic anoxia.

(a) *Shallow breathing.*—Anoxia may be present from toxic depression of the respiratory centre causing shallow breathing. The depth of the breathing is also restrained by the pain due to the associated pleurisy. The breathing may perhaps be *reflexly* accelerated by abnormal afferent impulses from the affected lung.

(b) *Unærated areas of lung.*—If the blood vessels supplying the consolidated parts are obstructed, all the blood must pass through the functioning parts of the lung and is there arterialized. But if the blood vessels of the pneumonic area are *patent*, the blood circulates through this region unchanged and leaves it in just as venous a state as that in which it arrived. The "arterial" blood is then a mixture of oxygenated and unoxygenated blood.

(c) *Altered alveolar permeability.*—If *œdema* of the lung is present, too, from heart failure, the condition is still further aggravated.

(iii) COLLAPSE OF LUNG.—If part of a lung is collapsed, the production of anoxia depends on the persistence of the local pulmonary blood flow together with the absence of local aeration. This most commonly occurs in collapse due to bronchial obstruction, *e.g.* owing to inhalation of a plug of mucus following an anæsthetic; the affected lobe or segment collapses as the air is absorbed distal to the block. Anoxia is much less common in cases of "extrinsic" collapse, *e.g.* hydro- or pneumothorax, because the compression is generally not sufficient to occlude the alveoli. More rarely the absence of anoxia is due to compression occluding both the lung capillaries and the alveoli.[2]

(iv) The results of *emphysema, tracheal and bronchial stenosis,* are considered on pp. 454, 457–459.

(3) ABNORMALITIES OF HEART AND BLOOD VESSELS.[3]—In certain forms of congenital heart disease, especially in the condition known as *Fallot's tetralogy*, a proportion of the venous blood is shunted directly through the heart into the arterial side of the circulation (p. 452); as a result, part of the blood reaching the systemic arteries has never passed through the lungs at all, and anoxia is present. Here again the "arterial" blood is a mixture of oxygenated and unoxygenated blood. The degree of anoxia depends on the magnitude of the shunt.

2. **Anæmic Type.**—This is less serious in its effects than the anoxic form. As the oxygen tension in the blood is normal, the rate of tissue oxidation is maintained at its usual level. No increase in the pulmonary ventilation occurs at first; *i.e.* the breathing does not respond readily to a decrease in the volume of oxygen in the arterial blood, so long as the appropriate tension is maintained. At rest, the prejudicial effect on the tissues is relatively

[1] Lundsgaard, *Medicine*, 1925, *4*, 345.
[2] Bjork and Salen, *J. thorac. Surg.*, 1950, *20*, 933.
[3] Prinzmetal, *J. clin. Investig.*, 1941, *20*, 705.

slight. Of the 19 c.c. of oxygen normally present in arterial blood, only 5 c.c. are used up under resting conditions. As a person with 50% hæmoglobin carries 9·5 c.c. oxygen per 100 c.c. blood, his resting requirements are readily satisfied. But his capacity to do work is greatly diminished because he has not the normal reserves of oxygen in the blood to call upon. (The anoxia of hæmorrhage is described on p. 81.)

(i) Lack of Hæmoglobin.—In severe anæmias (p. 196), the total hæmoglobin content is diminished, and the venous blood is very reduced. In exercise, the chief method which is available in these subjects for increasing the oxygen supplies to the tissues is to increase the cardiac output. Acceleration of the pulse occurs with slight exertion.

(ii) Altered Hæmoglobin.—In poisoning with the *nitrites*, nitric-oxide-hæmoglobin and some methæmoglobin are formed. With large doses of *chlorates* or *sulphonamide* derivatives, methæmoglobin is formed; marked cyanosis may be present, owing to the presence in the blood of this coffee-coloured compound. The patient may, however, feel well so long as enough normally functioning hæmoglobin (*i.e.* capable of transporting oxygen) is available, because the arterial oxygen tension is normal.[1] (cf. p. 454.)

(iii) Carbon Monoxide Poisoning.—This condition needs special consideration. It is a method commonly employed by intending suicides, and occurs accidentally from escapes of lighting gas, fires, and underground explosions.

CO acts by displacing oxygen from its combination with hæmoglobin, forming a comparatively stable compound which is useless for gas-carrying purposes; CO is so poisonous because we are dependent on functioning hæmoglobin for oxygen transport. A cockroach, which has no hæmoglobin in its blood, can be kept alive in an atmosphere containing 80% CO and 20% O_2. If a mouse is placed in a glass vessel and exposed to two atmospheres pressure of oxygen and one atmosphere of CO, almost all the hæmoglobin combines with CO, and yet the mouse remains normal while at rest, and only seems to tire easily when climbing up the sides of the jar. At the oxygen pressure of the experiment, about 4·2 c.c. O_2 are present in *free solution* in 100 c.c. of plasma; this amount is sufficient to supply the resting requirements of the body.[2]

The *affinity* of carbon monoxide for hæmoglobin is about 300 times as great as that of oxygen; *e.g.* if 14% oxygen and 0·047% CO are present in the alveoli, the hæmoglobin of the blood shares itself out equally between the two gases. [Minute concentrations of CO, if constantly maintained in the alveoli, can thus produce grave effects.] The symptoms then present are very serious. The subject cannot walk any distance, and sometimes can scarcely rise from his chair without collapsing on to the ground. When over 50% of the hæmoglobin is thrown out of action the slightest exertion may cause fainting. In fatal cases the saturation of the blood with CO is usually over 80%, but may be only 60%. The symptoms of CO poisoning are clearly those of oxygen lack. It is noteworthy that though dizziness, faintness, and mental disturbances are present because of the cerebral anoxia there is

[1] If an animal in this condition is kept in an atmosphere of oxygen for a sufficiently long period, the methæmoglobin gradually disappears from the blood and full recovery occurs. For discussion see Killick, *Physiol. Rev.*, 1940, *20*, 313.

[2] In much higher concentrations CO can directly arrest tissue oxidation.

no appreciable increase in breathing and no "air-hunger," because the oxygen tension in the blood is normal.

It is important to consider why the symptoms are so severe when 50% of the hæmoglobin is not functioning as a result of combination with CO, and why they are relatively so mild when the hæmoglobin is reduced to the same extent in anæmia. It is found that the presence of COHb *alters the dissociation curve of the functioning hæmoglobin* which is still present. The curve is shifted to the *left* and the S-shape disappears; in other words, hæmoglobin will now only give off oxygen in useful amounts when the oxygen tension is very low. COHb thus *prevents* adequate amounts of oxygen being given off to the tissues, which accordingly suffer very seriously. (cf. Fig. 249.)

If the patient is exposed to high pressures of oxygen (especially if CO_2 is added to the inspired air to stimulate breathing), CO is progressively displaced from its union with hæmoglobin by the mass action of the excess oxygen, and a normal state of the blood is gradually restored.

3. **Stagnant[1] Type.**—This occurs when the cardiac output and blood flow to the organs is diminished because of *heart failure* (p. 295), *impaired venous return*, *hæmorrhage*, or *shock*. The tension of oxygen in the blood is normal, but the amount reaching the tissues is inadequate. The rate of tissue oxidation is normal, because oxygen is supplied at a high pressure head. As the blood circulates more slowly in the tissues there is more time available for reduction of oxyhæmoglobin. Furthermore, the impaired circulation causes CO_2 accumulation in the tissues which facilitates the giving-off of oxygen. Thus the tissues make the most effective use of what oxygen does reach them in the blood.

4. **Histotoxic Type.**—This occurs in poisoning with *cyanide*, which interferes with tissue oxidation, by paralysing cytochrome oxidase (p. 854). *Narcotics* also depress tissue oxidation by interfering with dehydrogenase systems (p. 854).

The following Table (from Van Slyke) shows the circumstances under which the different varieties of anoxia may lower the oxygen tension in the venous blood (and therefore presumably in the tissues) to the same level and so produce comparable effects on the tissues:

	Arterial Blood.					Venous Blood.		
	O_2 Capacity.	O_2 Content.	Per Cent. Saturation Hb.	O_2 Tension mm.	O_2 Lost to Tissues.	O_2 Content.	Per Cent. Saturation Hb.	Calculated O_2 Tension.
Normal	20	19	95	90	4·2	14·8	74	40
Anoxic	20	14·8	74	43	4·2	10·6	53	28
Anæmic	10	9·5	95	90	4·2	5·3	53	28
CO poisoning	15·3[2]	14·7	74	90	4·2	10·1	50	28
Stagnant	20	19	95	90	8·4	10·6	53	28

[1] This might be better termed *ischæmic* type.

[2] The rest of the hæmoglobin is bound with CO. The figure underlined in each case is the cause of the anoxia.

CYANOSIS [1]

By cyanosis is meant a blue or bluish colour of the skin, mucous membranes, or the deeper organs. It may be general in distribution, but is usually more marked in certain regions; it is most often localized in the lips, nose, cheeks, ears, hands and feet. Cyanosis is due to a change in the character of the circulating blood, and so disappears if the blood is squeezed out from the part.

The following average standards are assumed in the subsequent discussion:

Normal arterial blood = 95% saturated with oxygen.
19 c.c.% O_2 content.
1 c.c.% O_2 *unsaturation* (A).
O_2 utilization by tissues = 5 c.c.%.
Mixed venous blood = 70% saturated with oxygen.
14 c.c.% O_2 content.
6 c.c.% O_2 *unsaturation* (V).

Cyanosis depends on the *absolute amount of reduced hæmoglobin present in the blood* (occasionally it is due to the presence of other dark hæmoglobin derivatives, such as met- or sulph-hæmoglobin, p. 454); the amount of oxygenated hæmoglobin present is of little importance. The amount of reduced hæmoglobin in blood can be expressed as volumes per cent. of oxygen *unsaturation*, *e.g.* 1 g-% reduced Hb = 1·34 c.c.% O_2 unsaturation = 18·66 c.c.% O_2 content. Average unsaturation in the capillary blood is taken as the mean of the arterial and venous unsaturation $= \frac{A+V}{2} = \frac{1+6}{2} = 3{\cdot}5$ c.c.% (= 16·5 c.c. O_2 content %). About 5 g-% reduced Hb or 6·7 c.c.% O_2 unsaturation is the least amount which must be present in the capillary blood to produce cyanosis; but individual clinicians vary in the ease with which they recognize the presence of slight degrees of cyanosis. For the reason stated above an anæmic individual who has less than 5 g. Hb% cannot usually become cyanosed.

The cyanotic colour is due to the blood in the *minute vessels*, *i.e.* capillaries, and possibly in the arterioles and venules of the subpapillary plexus as well.[2]

[1] Lundsgaard and Van Slyke, *Medicine*, 1923, 2, 1.

[2] Factors modifying Cyanosis.—(i) Thickness of the epidermis, which itself has no vessels. (ii) Presence of normal or pathological pigments in the skin. (iii) Variations in the colour of the plasma from varying concentration of lipoids, white cells in leukæmia or pigments. (iv) Variation in the concentration of oxyhæmoglobin in the blood. (v) Variation, in the number, width, and length of the blood-filled capillaries in a given area.

These modifying factors cannot alone produce cyanosis. They can influence the amount of reduced hæmoglobin (O_2 unsaturation) which is necessary to give the skin a perceptibly blue colour, and can modify the exact shade of colour produced.

Factor (v) needs further consideration: (*a*) With the skin microscope, variations are noted in the deneness of the capillary network in different parts of the skin; great variations are present in the capillary distribution in different individuals. (*b*) The influence of *temperature* is important. When the surrounding atmosphere becomes colder the arterioles are constricted and the capillaries dilated. The blood flow becomes slow, and extensive reduction of the blood may take place (cf. p. 328). (*c*) An *increased venous pressure* tends to increase the number, width, and length of the blood-filled capillaries, and the deeper vessels become visible; the production of cyanosis is thus facilitated. (*d*) Increase in the *blood volume* may have a similar effect. These facts would help to explain the ready production of cyanosis in *polycythæmia*, in which there is an increased blood volume and possibly a hyperplastic condition of the capillaries as well.

Causes of Cyanosis.—It is obvious that the causes of anoxia and cyanosis must, to a large extent, overlap. The following classification (after Lundsgaard and Van Slyke) may be used:

1. *All the blood passes through lung tissue accessible to air, but for various reasons oxygenation of the blood in the lungs is incomplete.*

This occurs in—(i) Low O_2 pressure in the alveoli, *e.g.* at high altitudes, breathing oxygen-poor mixtures, or from restricted pulmonary ventilation (from obstruction to the respiratory passages or from failure of the respiratory mechanisms).

(ii) Swelling and thickening of alveolar wall and consequently diminished permeability to oxygen (" pneumonosis ").

(iii) Inefficient alveolar ventilation from shallow breathing, or in emphysema.

If the arterial oxygen content falls to 16 c.c.%, cyanosis is present.

$$A = 4 \text{ c.c.}; \; V = 9; \; \frac{A+V}{2} = 6{\cdot}5$$

which is the threshold for cyanosis.

2. *Presence of an unærated shunt between the veins and arteries.*

(i) In certain forms of congenital heart disease (*infra*).

(ii) Collapse or consolidation of part of a lung with persistence of the blood flow through the region. It can be calculated that if *over one-third* of the output of the heart is shunted from the venous to the arterial side without being ærated, cyanosis develops.

Group 1 is relieved by the administration of oxygen, which raises the oxygen pressure in the alveoli and (in 1 (ii)) increases the rate of diffusion into the blood.[1] Group 2 (i) is completely unaffected by oxygen therapy; 2 (ii) can only be helped if the pulmonary collapse is not quite complete, in which case oxygen inhalation may enable some oxygen to get into the almost functionless areas.

3. *Greater reduction of oxyhæmoglobin in the tissues.*

(i) Local chilling of a part (within certain limits, p. 328), or rise of venous pressure which retards the circulation, will slow down the local rate of blood flow and so increase the reduction of the blood.

(ii) Diminished blood flow to the tissues in heart failure is accompanied, as noted (p. 450), by high oxygen utilization in the tissues. Cyanosis may not be present at rest, but the increased oxygen consumption which accompanies activity may be sufficient to raise the reduced hæmoglobin content to the required level.

The oxygen consumption of the tissues must rise to 12 c.c. per 100 c.c. of blood flow (normal = 5 c.c.), *i.e.* to two and a half times the normal, before cyanosis can be produced by this factor alone.

Certain clinical conditions will be more fully considered.

(1) CONGENITAL HEART DISEASE.[2]—Some forms of congenital heart disease are *not* associated with cyanosis, *e.g. patent ductus arteriosus* (p. 383) or *septal defects* in which the flow is from the left side to the right side of the heart. In both these conditions the pulmonary blood flow is actually

[1] Inhalation of 100% O_2 in patients with chronic anoxia, *e.g.* emphysema, may lead to mental changes, coma, and death; the cerebrospinal fluid pressure is high. The mechanism involved is obscure (Comroe *et al.*, *J. Amer. med. Assoc.*, 1950, *143*, 1044).

[2] Wood, *Brit. med. J.*, 1950, ii. 639, 693.

increased and the venous blood is fully oxygenated in the lungs. In *coarctation of the aorta* the obstruction is in the aortic arch and is, therefore, on the systemic side of the circulation; in this condition cyanosis may develop from peripheral causes in the distal parts of the body which are inadequately supplied with blood.

Cyanosis of *central* origin is present when there is a *shunt from the venous to the arterial side* of the circulation. This occurs most commonly in *Fallot's tetralogy* (see below). Cyanosis is rare in *simple pulmonary stenosis* unless this anomaly is also associated with a shunt from the right to the left side of the heart.

Peripheral factors may as usual increase the tendency to cyanosis, *e.g. exercise*, which raises the oxygen utilization in the capillaries, or external *cold*.

Polycythæmia, *i.e.* an increase in the number of red cells per c.mm. to 7 or 8 million, occurs when congenital heart disease is associated with chronic anoxia. As it increases the viscosity of the blood the rate of flow through the capillaries is slowed; the capillaries are increased in width and number.

Fallot's Tetralogy.—The abnormalities which constitute this syndrome are: (i) a small, narrow pulmonary artery; (ii) a patent interventricular septum; (iii) relative displacement of the ascending arch of the aorta to the right so that its mouth lies above the patent part of the interventricular septum; it consequently receives blood from both ventricles; (iv) the circulation is carried on mainly by the right ventricle which is greatly hypertrophied; the left ventricle is atrophic. A considerable proportion of the right ventricular outflow is pumped directly into the aorta. The intensity of the cyanosis depends on the magnitude of this shunt of venous blood into the systemic circulation. Some of the patients are greatly helped by the following operation: one of the subclavian arteries is anastomosed to a suitable point on the pulmonary artery.[1] An artificial patent ductus arteriosus is thus established whch diverts some of the partially oxygenated aortic blood through the lungs. The results in a successful case are as follows: (i) a fraction of the venous return continues to pass from the right ventricle into the lungs through the narrow pulmonary artery; (ii) in addition some of the aortic blood (which is, in part, venous blood derived from the right ventricle) is oxygenated further; this blood returns to the left heart and is again ejected into the aorta; (iii) the oxygen content of the arterial blood reaching the tissues consequently rises, sometimes markedly. In one group of patients the average arterial oxygen saturation rose after the operation from 49% to 76% (under non-basal conditions). The polycythæmia and the cyanosis decrease and the exercise tolerance is improved.

(2) Mitral Lesions.—There is increased left auricular and pulmonary capillary pressure; the resulting pulmonary œdema or swelling of the alveolar wall decreases oxygenation of blood in the lungs.

(3) Heart Failure.—When heart failure develops from any cause, cyanosis is common. Œdema of the lung, hydrothorax, shallow breathing, all result in under-arterialization of the blood in the lungs. The diminished blood flow to the tissues produces capillary stagnation, high oxygen utilization and greater reduction of the blood, and therefore cyanosis (cf. p. 450).

[1] Blalock, *Bull. N.Y. Acad. Med.*, 1946, 22, 57.

(4) Tracheal and Bronchial Stenosis.—In these conditions a normal degree of contraction of the inspiratory muscles and enlargement of the thoracic cavity leads to a subnormal intake of tidal air past the obstruction into the alveoli. Greater respiratory efforts are consequently made; but if the obstruction is sufficiently severe, even maximal respiratory efforts (producing great subjective distress) do not produce adequate alveolar ventilation. The alveolar O_2 tension falls because the rate of passage of O_2 from the lungs into the arterial blood is greater than the rate at which fresh supplies of O_2 are obtained from the outside air. CO_2 accumulation occurs for the same reasons. *Asphyxia* (*i.e.* O_2 lack plus CO_2 excess) is thus present. In tracheal stenosis, arterial anoxia, dyspnœa, and cyanosis are often well marked; in bronchial stenosis, however, cyanosis is uncommon. At least one-third of the blood must pass through the unærated channel (which means that at least two-thirds of one lung must be shut off as a result of bronchial obstruction) before cyanosis can appear from this cause (cf. p. 452).

(5) In Asthma there is deficient pulmonary ventilation from bronchiolar obstruction. In Bronchiolitis there is patchy collapse of parts of the lung. The cyanosis of Emphysema is often very marked. It is in part due to inadequate ventilation of those alveoli through which the pulmonary circulation passes; there is therefore a lowered oxygen saturation of the arterial blood and resulting cyanosis which is more marked during exertion or after the development of heart failure (for full discussion see p. 459).

In Lobar Pneumonia and Œdema of the Lung there is arterial anoxia for reasons sufficiently explained (pp. 447, 448).

Clinically, *acute pneumothorax* is not usually associated with cyanosis; if the lung collapse is small cyanosis would not be expected; if the lung collapse is severe, the vessels are usually occluded and the main symptoms are *pain* and *dyspnœa* which need urgent treatment. In *chronic pneumothorax*, cyanosis occurs if the blood flow in the affected lung is not paralleled by the local alveolar ventilation. The opposite lung is protected to some extent from the effects of the raised intrapleural pressure by the presence of pathologically increased rigidity of the mediastinal structures (p. 370).

(6) Alteration in the Hæmoglobin of the Blood.—(i) *Methæmoglobinæmia.*—Certain drugs and poisons, *e.g.* chlorates or sulphonamides, lead to the formation of *methæmoglobin*, which gives the blood a dark colour (p. 174). Some of the coal-tar preparations—acetanilide, sulphonal, trional—produce a similar result. Methæmoglobinæmia may also occur as a *congenital* disorder of unknown origin. The administration of large doses of *ascorbic acid* (*e.g.* 300–600 mg. daily) in these patients may lead to the conversion of much of the methæmoglobin into oxyhæmoglobin.

(ii) *Sulph-hæmoglobinæmia.*—In this rare condition sulph-hæmoglobin is formed, giving rise to cyanosis of a leaden hue. The pigment is wholly intracorpuscular; its mode of formation is not yet clearly understood. It must be remembered that sulph-hæmoglobin is not formed in blood when excess of H_2S is present in cases of poisoning with this substance. A strong *reducing agent* of unknown composition is present in the blood, urine, and saliva in sulph-hæmoglobinæmia. If the serum of these patients is added to whole blood, reduced hæmoglobin is soon formed. It has been shown that powerful reducing agents permit the formation of sulph-hæmoglobin from oxyhæmoglobin in the presence of minute traces of H_2S. It is suggested that

the reducing agent found in the blood of these patients allows the normal traces of H_2S which are absorbed into the blood from the bowel to effect this conversion. A bacillus has been isolated from patients which bears some ill-understood relationship to the disease. It is readily recognizable, as it can disintegrate amino-acids to form nitrites (" nitroso bacillus ").

DYSPNŒA[1]

Dyspnœa—Normally, breathing goes on without intruding on consciousness. Dyspnœa literally means difficult breathing. Meakins defines it as " consciousness of the necessity for increased respiratory effort." When the breathing enters consciousness unpleasantly and produces discomfort, it is called dyspnœa. This definition is not entirely satisfactory because it excludes the grave disturbances of breathing which may occur in unconscious subjects (*e.g.* in diabetic or uræmic coma). *Hyperpnœa* simply means increased breathing; for a time it does not impinge on consciousness, and so represents a stage preceding the onset of dyspnœa. An ordinary person is not aware of any increase in the breathing until the pulmonary ventilation is doubled. Real discomfort develops when the ventilation is increased four- or fivefold; this level of ventilation is called the *dyspnœa point.* Dyspnœa is not wholly a pathological phenomenon, for, as noted (p. 438), dyspnœa develops in normal subjects during strenuous exertion.

The following factors require further consideration :—

(1) Breathing Reserve (BR).—The *maximal breathing capacity* (MBC) is the maximal voluntary pulmonary ventilation in L per minute, determined during a 15 sec. period. The pulmonary ventilation in L per minute at rest or under any other specified conditions is designated PV. The *breathing reserve* (BR) under the specified conditions is MBC$-$PV. The *percentage breathing reserve* (%BR) is (MBC$-$PV)/(MBC)$\times$100. It is also called the *dyspnœic index.* If its value falls below 60% (range 60-70%) dyspnœa is generally present. The %BR may be lowered owing to a decrease in maximal breathing capacity (MBC) or a rise in pulmonary ventilation (PV).

Example: Normal resting person, MBC=100; PV=8; %BR=92%. When owing to exertion PV increases to 40L per minute the %BR falls to 60% and dyspnœa is present.

(2) Vital Capacity.—A decrease in vital capacity decreases the maximal breathing capacity and thus the percentage breathing reserve; it therefore predisposes to dyspnœa. As the depth of breathing approaches the vital capacity the sense of discomfort increases.

(3) Mechanical Efficiency.—A person with a low mechanical efficiency uses more energy than a normal person to do a given amount of work: his O_2 consumption and pulmonary ventilation are correspondingly greater and he thus develops dyspnœa earlier.

Pathological Dyspnœa.—The causes of dyspnœa may be classified as follows:

[1] Means, *Medicine,* 1924, *3,* 309–416. Haldane and Priestley, *Respiration,* new edn., Oxford, 1935. Harrison, *Failure of the Circulation,* 2nd edn., Baltimore, 1939; Christie, *Quart. J. Med.,* 1938, *7,* 115. Cournand *et al., Medicine,* 1948, *27,* 243; 1949, *28,* 1, 201.

1. **From Increased Metabolism.**—Patients with *exophthalmic goitre* have an abnormally high rate of metabolism at rest, and therefore their resting pulmonary ventilation may be 50–100% above normal, depending on the severity of the case. The vital capacity is reduced in this condition, and the mechanical efficiency is lower than normal, *e.g.* only 15%. No dyspnœa is present as a rule at rest; but when doing work the patient is handicapped in several ways: (i) he commences with a high resting pulmonary ventilation; (ii) being an inefficient machine, his ventilation increases very steeply, as he has to liberate an excessive amount of energy to perform any task; (iii) owing to the low vital capacity, subjective symptoms arise when the hyperpnœa is yet moderate; (iv) in some cases heart failure is an additional factor to be considered (p. 380).

2. **From Metabolic Acidæmia** (cf. p. 99).—(1) Some cases of *nephritis* show an acidæmia, owing to the failure of the kidney to excrete adequately acid radicals, chiefly phosphate. Breathing is therefore stimulated, and the ventilation may reach the dyspnœa point. The increased breathing helps to restore the normal blood reaction by washing out CO_2; the CO_2 tension in alveolar air and arterial blood is consequently very low. In one uræmic patient with a blood-urea concentration of 332 mg-% the pulmonary ventilation was 51 litres, the alveolar CO_2 6·4 mm. Hg, and the total CO_2 content of the blood only 12 c.c.%. The administration of alkali in adequate amounts temporarily restores the normal *p*H of the blood and abolishes the hyperpnœa or dyspnœa and the coma which may have been present. There is not, however, a constant ratio in nephritis between the degree of acidæmia and the pulmonary ventilation, as the *sensitivity* of the respiratory centre may be depressed by the disease (cf. pp. 76, 77).

(2) *Ingestion* of substances which give rise to acids in the body, *e.g.* methyl alcohol (which is converted into formic acid) and ammonium chloride (which releases H^+ ions) (p. 397), may produce dyspnœa.

(3) *Diabetes Mellitus.*—In severe diabetes, acetoacetic and β-hydroxybutyric acids are formed in excessive amounts, pass into the blood and are neutralized by the plasma bicarbonate. Compensatory hyperpnœa develops which lowers the alveolar and arterial CO_2 tension (p. 102). The breathing is slow and deep but easy, and the pulmonary ventilation is rarely increased fivefold, *i.e.* to the dyspnœa point. When the breathing is truly laboured the patient is usually comatose so that no "subjective" dyspnœa is present (cf. pp. 102, 925).

3. **From Oxygen Lack.**—Anoxic stimulation of breathing has already been fully discussed (p. 398). *Anæmic* and *stagnant* anoxia (except severe hæmorrhage) tend to produce little increase in breathing at rest as the oxygen tension in the arterial blood is normal. In the *anoxic* form the respiratory centre is stimulated reflexly by the subnormal oxygen tension in the blood; and if adequate time is allowed for compensation to be established, the breathing may be very greatly increased (p. 400).

4. **From Mechanical and Nervous Hindrance to the Respiratory Movements.**—In this group the main factor predisposing to dyspnœa is generally the decrease in maximal breathing capacity (MBC) and in vital capacity. In severe chronic pulmonary disease, MBC may fall to 20–30% of normal. Dyspnœa is brought on or aggravated when the pulmonary ventilation is increased from any cause. The fall in percentage breathing

reserve (p. 455) is well-correlated with the appearance and extent of the dyspnœa.

(1) *Bronchial Obstruction and other Conditions associated with Pulmonary Collapse or Consolidation.*—(i) If one *bronchus is blocked* breathing is stimulated; the degree of dyspnœa is directly related to the speed and completeness of the collapse of the lung. Experimentally, the hyperpnœa often disappears on cutting the vagus on the affected side, indicating that it was *reflexly* produced. Anoxia may be an additional contributory factor (cf. pp. 448, 454).

(ii) In *acute* closed pneumothorax due to the introduction of a small volume of air into the pleural cavity, the respiratory rate is increased, but returns to normal after bilateral vagal section (Fig. 266). [In *acute large open* pneumothorax the respiratory changes are different (p. 368).] These observations suggest that in *acute pulmonary collapse* or *consolidation* from any cause an *excitatory* vagal reflex may be an important factor in producing increased pulmonary ventilation, quite independently of any changes in the chemical composition of the arterial blood.

(iii) In more *chronic* conditions associated with collapse of the lung this vagal reflex effect may not be in evidence; thus in cases of *chronic closed*

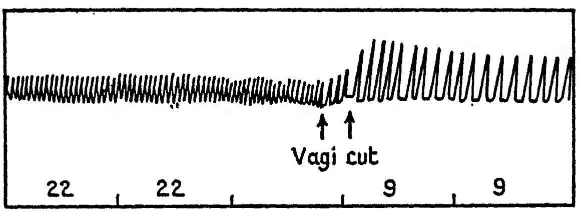


Fig. 266.—Reflex Effect of Closed Pneumothorax on Respiration. (After Harrison, *Failure of the Circulation.*)

Dog. Record of respiration. Vagi intact. 200 c.c. of air introduced into pleural cavity 13 minutes previously. Note rapid shallow breathing. At arrow, cut both vagi. Breathing becomes slow and deep. The figures on the abscissa show the rate of breathing per minute.

pneumothorax, for example, the breathing at rest is quite comfortable. In response to exercise or to excess of CO_2 in the inspired air, the breathing increases in a normal manner; dyspnœa, however, is present when the ventilation has increased only threefold owing to the associated decrease in maximal breathing capacity.

(2) *Asthma.*—In asthma, bronchial constriction is present which diminishes the vital capacity and prevents the easy passage of air into and (and even more) out of the lungs (p. 409). During an acute attack, the breathing capacity is so limited by the bronchial spasm, and the effort needed to displace air from the lungs is so great, that intense dyspnœa is present; the onset of *asphyxia* stimulates the respiratory efforts still further.

(3) *Paralysis of Respiratory Muscles.*—Paralysis of the *diaphragm*, for example, or interference with its action by great abdominal distension produces respiratory embarrassment especially when it is of sudden occurrence. Here again the breathing capacity is diminished. In severe cases the asphyxial element comes into operation and soon becomes the dominating factor.

(4) *Obstruction to Main Air-way.*—When the resistance is *slight*, the breathing becomes slower and deeper. If the resistance is *excessive* and continues for a long time, the breathing ultimately becomes increasingly shallow and more frequent. The alveolar ventilation becomes inadequate in consequence, and arterial anoxia develops; if the obstruction is still more severe CO_2 retention also occurs. These points must be borne in mind when considering the effects of wearing a respirator.

When the air-way is free, the hyperpnœa of exercise can be maintained for a long time, but when the breathing is obstructed, the "fatigued" respiratory centre cannot keep up the hyperpnœa of work in the normal manner.

(5) *Effects of Disturbed Pulmonary Circulation.*—(i) Multiple *embolism* of the pulmonary arterioles can be produced by intravenous injection of starch granules. With a moderate degree of embolism insufficient to produce

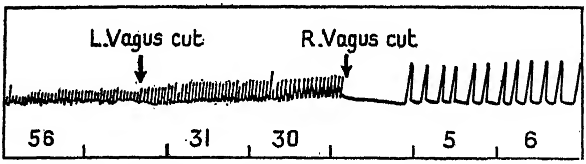


FIG. 267.—Reflex Effect of Pulmonary Congestion on Respiration. (After Harrison, *Failure of the Circulation*, 2nd edn., 1939.)

Dog. Record of respiration. Vagi intact. Pulmonary vessels congested by injecting fluid into pulmonary artery on one side after ligature of pulmonary vein. Note rapid shallow breathing. At first arrow, cut left vagus; at second arrow, cut right vagus. Rate of breathing per minute is shown by figures on base line. Breathing becomes slower and deeper.

death, there is a great increase in the frequency and a diminished depth of respiration to half or one-quarter normal. If the vagi are now divided slow respiration of normal depth develops. The nerve endings involved have not been identified but are probably vascular stretch receptors in the distal part of the pulmonary vascular bed.[1]

(ii) If engorgement of the vessels in one lung is produced by obstructing the pulmonary vein, breathing becomes much more rapid; the effect is reflexly produced and is abolished by cutting the vagus nerve on the affected side (Fig. 267). Clinically, *pulmonary congestion* is associated with increased pulmonary ventilation, which is probably produced in the same way.

In considering the production of dyspnœa in (1)–(5) above, the principles enunciated on p. 456 should be borne in mind.

Suppose that owing to pulmonary disease or other cause the maximal breathing capacity is reduced to 50 L; if the resting pulmonary ventilation is 8 L, the %BR is $(50-8)/50\times100=84\%$, which is above the level for dyspnœa. If the pulmonary ventilation is increased by exertion, reflexly, or by chemical factors (O_2 lack, CO_2 excess or asphyxia) to 15 L the %BR is $(50-15)/50\times100=60\%$, and dyspnœa is present.

Suppose that the MBC is reduced to 30 L. An increase in the pulmonary ventilation to 12 L would lower the %BR to 60% and cause dyspnœa. If the MBC is reduced to 20 L, dyspnœa will be present at rest if the pulmonary ventilation is 8 L.

[1] Whitteridge, *Physiol. Rev.*, 1950, *30*, 475.

(6) *Emphysema.*[1]—The disease is discussed for convenience at this point. The chief *structural* changes are as follows: The elastic framework of the alveoli is probably weakened. The walls of the air-sacs are stretched, thin, and ruptured at many points, resulting in the formation of large irregular spaces. Many pulmonary capillaries running in the walls of the most distended air-sacs are obliterated. The respiratory bronchioles are said to be dilated. Post-mortem the lungs do not collapse readily (because of their loss of elasticity); the free margins are thinned out and show large bulbous cavities.

Clinically, important hæmo-respiratory changes are noted:

(i) The *intrapleural pressure* is more nearly atmospheric than normal, indicating lessened elastic recoil of the lungs (p. 368) and offering further evidence of damage to the pulmonary elastic tissue.

(ii) Alterations take place in the *subdivisions* of the *lung volume* (Fig. 268):

(*a*) The residual air is increased, mainly because the patient cannot empty his lungs by means of a full expiration as completely as normal.

(*b*) The reserve air is decreased, but the resting lung volume at the end of a normal expiration (functional residual air) is generally increased.

(*c*) The vital capacity is reduced.

(*d*) In a mild case the tidal air and the total pulmonary ventilation are fairly normal.

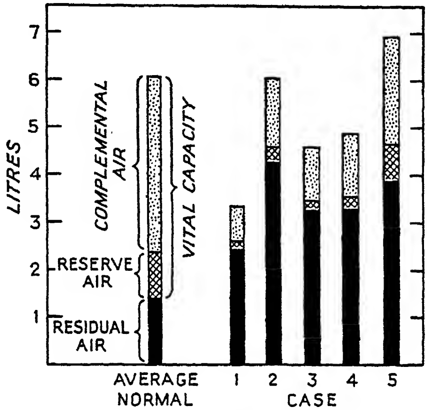


FIG. 268.—Chart showing the Lung Volume and its Subdivisions in the Normal and in Emphysema. Cases are arranged from left to right in order of severity of the dyspnœa. (Christie, *Brit. med. J.*, 1944, i, 105.)

The reduction of vital capacity will obviously impair the response to exercise and induce dyspnœa more readily. But as the resting pulmonary ventilation is about normal one would have expected that the gaseous composition of the blood would be normal; this, however, is not the case (*infra*).

(iii) (*a*) The arterial blood usually shows decreased saturation of the hæmoglobin with oxygen, indicative of a lowered arterial O_2 pressure. In a mild case the saturation may fall to 90%, in an advanced case to 80% or less. Oxygen therapy may be dangerous (p. 452).

(*b*) The CO_2 content and pressure in the arterial blood are raised; the latter may be 70–80 mm. Hg.

(*c*) There is no evidence of decreased permeability of the alveolar epithelia to O_2 and CO_2. The impaired arterialization of the blood must presumably be due to ineffective ventilation of the alveoli.

(*d*) The cardiac output (in the absence of heart failure) is unchanged and so the total pulmonary blood flow is undiminished.

(*e*) But, as has already been mentioned, many capillaries are obliterated, especially in the walls of the more dilated air-sacs. Atmospheric air which

[1] Christie, *Brit. med. J.*, 1944, i, 105.

reaches non-vascular air-sacs is wasted as far as arterializing the blood is concerned. Only a fraction of the tidal air reaches vascularized functioning air-sacs, which are consequently inadequately ventilated in relation to their blood flow : the arterial blood is incompletely oxygenated and CO_2 is retained. Inhalation of pure O_2 relieves the anoxia but does not relieve the CO_2 retention and may be dangerous (p. 452). There is thus a state of anoxic anoxia with increased liability to *cyanosis* (p. 454) and acidæmia due to CO_2 retention; both these conditions, especially the latter, tend to produce dyspnœa.

(iv) During exertion, breathing is stimulated by increased CO_2 formation and retention, and by an aggravation of the anoxia. As the vital capacity is reduced, a moderate degree of hyperpnœa produces considerable subjective discomfort, so that dyspnœa readily develops. If cardiac failure occurs (because of the overwork of the right ventricle in overcoming the greater pulmonary resistance) the dyspnœa is aggravated.

(7) *Neurogenic Dyspnœa.*—In certain organic diseases of the nervous system, especially if they involve the region of the hypothalamus (p. 716), deep or panting breathing at rest may be present, with the usual complications of overventilation at rest (Fig. 56).

Shallow rapid breathing also occurs in neurasthenia and other functional nervous disorders.

5. **Cardiac Dyspnœa.**[1]—In mild cases of circulatory failure, dyspnœa only develops with muscular effort, but effort of an order which would not cause distress in a normal person. When failure is more marked dyspnœa also occurs at rest. (i) It may occur when the patient is lying down, but not when he is sitting up (*orthopnœa*). (ii) It may be absent or mild in the morning, develops gradually during the day and becomes worse in the evening (*evening dyspnœa*). (iii) It may consist of paroxysms of dyspnœa, waking the patient up at night (*paroxysmal dyspnœa, cardiac asthma*). (iv) It may be periodic in character (*Cheyne-Stokes respiration*). (v) In extreme cases it may be present throughout the day, getting worse at night, and only relieved by opiates (*continuous dyspnœa*). (vi) Finally the dyspnœa may be aggravated by the development of pulmonary complications such as hydrothorax, infarction or pneumonia.

(1) **Simple Cardiac Failure.**—Let us assume that the primary disturbance affects the *left ventricle, e.g.* because of *aortic incompetence* or *hypertension.* For a time the left ventricle responds to the extra load; contracts more vigorously and maintains an efficient circulation. As the left ventricle fails it begins to dilate, *i.e.* a certain volume of residual blood remains in it at the end of systole. This leads to interference with the emptying of the left auricle and a rise of pressure within it. If the primary disturbance is *mitral stenosis,* left auricular dilatation develops and as failure occurs, the pressure within it progressively rises. The venous return from the lungs is consequently hampered, leading to *pulmonary congestion.* This in turn diminishes the vital capacity and so lowers the dyspnœa point; it also reflexly stimulates breathing (p. 458). Gradually the right ventricle fails because its work is increased owing to the greater pressure in the pulmonary circuit. This in turn leads to right ventricular dilatation, rise in right intraventricular pressure (Fig. 150),

[1] Fraser, *Lancet,* 1927, i, 429 *et seq.* Christie, *Quart. J. Med.,* 1938, 7, 115. Harrison, *Failure of the Circulation,* 2nd edn., Baltimore, 1939, pp. 186–313.

interference with the emptying of the right auricle, rise of right intra-auricular pressure and then an *increase of systemic venous pressure.* This rise of venous pressure reflexly further stimulates respiration.

The patient with a failing circulation thus tends to have a higher resting pulmonary ventilation owing to the reflex stimulation set up from the congested systemic veins and pulmonary vessels. In cases of heart disease in which an adequate circulation is maintained, the metabolic rate is normal; but when heart failure develops, the metabolism for some unknown reason is often raised by 25–50% above normal. This, of course, means that the pulmonary ventilation at rest must be correspondingly increased.

Effects of Exertion.—During *exertion* the muscular movements reflexly stimulate breathing further; the venous return increases, but as the heart cannot cope adequately with the increased load the venous and pulmonary congestion is intensified. As a result breathing is stimulated, *not only by the normal chemical changes* in the blood produced by exercise, but *reflexly also.* Consequently, the increase in pulmonary ventilation for a given degree of work is greater than in a normal person, and the breathing returns to normal more slowly.

If the exertion is more severe, all the factors enumerated above which stimulate breathing operate more powerfully. The heart is still more incompetent to deal with the greater venous return and its *output does not rise* to the normal degree. The insufficient blood flow through the lungs leads to excessive *CO_2 retention* and an *inadequate O_2 uptake.* The resulting anoxia and CO_2 accumulation further stimulate breathing. The blood flow to the active muscles does not increase in proportion to their needs, so that lactic acid accumulates progressively. It is also claimed that the buffering power of the muscles is depressed for some unknown reason so that a larger proportion of the lactic acid passes out into the blood, leading to *acidæmia* and further increase in breathing. Throughout it must be remembered that the *dyspnœa point is lowered* owing to diminished vital capacity, the mechanical efficiency is subnormal, and the *dyspnœa index* is lowered (p. 455).

Some of these considerations are summarized in Fig 269.

Other Contributory Factors.—The *vital capacity* and the *breathing reserve* are further diminished in heart disease by complications like hydrothorax. When the vital capacity is 70–90% of the normal, the patient is still able to do work, but dyspnœa develops on exertion; when it is 40–70%, very little work can be done; when it is below 40%, the patient is usually bedridden, with marked signs of cardiac insufficiency.

If œdema of the lung develops or if emphysema is a complicating factor, anoxia of the anoxic type develops owing to imperfect saturation of venous blood with oxygen as it passes through the lungs.

In some cases breathing is stimulated reflexly by the various factors to such an extent that the alveolar and arterial CO_2 tensions fall and *alkalœmia* of the arterial blood results (p. 101). In other cases impaired elimination of CO_2 may be the dominating factor with resulting acidæmia even at rest. Secondary renal failure may complicate the issue with further retention of non-volatile acids (p. 76).

(2) Orthopnœa.—In these patients it must be supposed that owing to venous and pulmonary engorgement and reduced vital capacity the breathing

Overwork or Disease of Left Ventricle → Dilatation of left ventricle.
↓
Mitral Stenosis → Rise of pressure in left auricle.
↓
Congestion of lungs → { *Decreased vital capacity → Lower dyspnœa point.* / *Reflexly stimulates breathing.* }
↓
Overwork or Disease of Right Ventricle → Dilatationof right ventricle.
↓
Tricuspid Stenosis → Rise of pressure in right auricle
↓
Increased venous pressure

(Rise of pressure in right auricle; Increased venous pressure) → *Reflexly stimulates breathing.*

Mild Exertion → (i) Further increase in venous return
(ii) Muscular movements
(iii) Increased CO_2 production

((i)–(iii)) → *Further stimulation of breathing.*

More Severe Exertion → (i–iii) Intensified →
↓
Inadequate Increase in Cardiac Output → CO_2 retention →
Inadequate Increase in Cardiac Output → Inadequate O_2 intake → anoxia →
↓
Inadequate Blood-flow to Muscles
↓
Increased Lactic Acid Formation
↓
Impaired Buffering Power of Muscles

(Increased Lactic Acid Formation; Impaired Buffering Power of Muscles) → More lactic acid in blood → acidæmia →

(all four arrows above) → *Still greater stimulation of breathing.*

FIG. 269.—Diagrammatic Representation of Causes of Cardiac Dyspnœa on Exertion. (Modified from Harrison, *Failure of the Circulation*, 2nd. edn., Baltimore, 1939.)

is initially close to the abnormally low dyspnœa point. Under such circumstances any factor further stimulating breathing induces actual dyspnœa. When the patient moves from the upright to the recumbent position there is a shift of blood from the abdomen to the thorax, with greater congestion of both the right auricle and the lungs; the breathing is consequently further stimulated and the vital capacity is further reduced (perhaps by as much as 25%). In addition, the associated high position of the diaphragm mechanically embarrasses breathing as a greater effort is now needed to move the enlarged liver and the other abdominal organs. Dyspnœa consequently develops rapidly in the recumbent position (orthopnœa). At a later stage dyspnœa may occur even in the sitting position.

(3) Paroxysmal Dyspnœa.—The attack generally sets in at night; the patient wakes up with a feeling of anxiety and oppression in the chest. He sits up, or even seeks the window for more air; his heart is pounding, he starts to cough and soon he is suffocating and fighting for breath. The dyspnœa is very distressing; the breathing is noisy and there may be bronchial spasm. There are signs of raised pressure in the pulmonary bed, *e.g.* palpable systolic pulsation over the right ventricular conus and pulmonary artery, succeeded by a distinct shock as the pulmonary valves close. The cough is unproductive at first, but soon becomes associated in severe attacks with a little frothy sputum, which is possibly blood-stained. In grave cases bubbling noises can be heard over the whole chest owing to the development of severe pulmonary œdema and there is a profuse expectoration of pink frothy fluid.

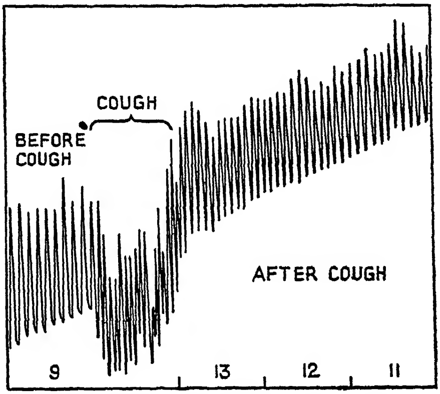


Fig. 270.—Effect of Bout of Coughing on Breathing. (Harrison, *Failure of the Circulation*, 2nd edn., 1939.)

Patient with congestive heart failure. Record of breathing. Voluntary bout of coughing for 1 minute. This is followed by more rapid breathing. Figures on abscissa represent rate of breathing per minute.

The mechanism of production of the attack has been much discussed. It is usually associated with *pulmonary congestion* and *raised pulmonary venous and capillary blood pressure*. In some patients the sequence of events may be as follows: The patient is the subject of some circulatory disorder that predisposes to the development of left ventricular failure and of dyspnœa (*e.g.* arterial hypertension, coronary artery disease or aortic disease). In such patients a short bout of voluntary coughing induces rapid breathing for some time (Fig. 270) which is attributed to pulmonary congestion. Even a voluntary deep breath may induce coughing and so give rise to an attack of dyspnœa. Severe coughing, by obstructing the venous return from the lungs, may lead to such intense pulmonary congestion that *œdema of the lungs* develops. The following explanation is offered for the frequency of occurrence of the attacks at night. One may suppose that at night, while the

patient is asleep, a good deal of mucus may collect in the tracheo-bronchial tract especially in a somewhat bronchitic patient; coughing develops, and the patient is woken up. The respiratory centre simultaneously "awakes" from its previous depressed state (cf. p. 405), its reflex excitability increases and the attack of coughing becomes more severe and induces pulmonary congestion and dyspnœa; the attack may progress to pulmonary œdema. Any other factor developing at night which reflexly stimulates breathing may induce an attack. A dose of morphine depresses the respiratory centre and so prevents the onset of dyspnœa.

In other cases the principal causal factor may be the actual development (for various reasons such as decreased coronary blood flow) of acute left ventricular failure, while the *right ventricle for a time goes on functioning normally*. The impaired efficiency of the left ventricle leads to congestion of the left auricle and consequently of the lungs, which is aggravated by the fact that the right ventricle initially discharges all the blood it receives in a normal way into the lungs. The congestion leads to reflex stimulation of breathing and pulmonary œdema; the dyspnœa is aggravated by chemical changes in the blood resulting from bronchial spasm or pulmonary œdema (cf. p. 117).

6. **Cheyne-Stokes (Periodic) Breathing.**—Periodic breathing occurs in a variety of disease conditions and is usually a symptom of considerable gravity; it is probably not always produced in exactly the same way.

(1) Experimentally, if the *intracranial pressure is raised*, no effect on respiration is observed till the pressure of the compressing fluid is almost equal to that in the carotid artery; at this stage there may be an increase in depth and frequency of respiration. If the intracranial pressure is raised further, apnœa develops from intense anæmia of the respiratory centre; this is often followed by periodic variations in both blood pressure and respiration. The sequence of events is probably as follows: medullary anæmia stimulates the vasomotor centre and the blood pressure gradually rises; as the medullary blood supply is restored the respiratory centre revives and breathing restarts either as a few uniform breaths, as a dwindling series, or as a group of breaths which rise to a maximum and then dwindle. Owing to the improved local circulation, the stimulus to the vasomotor centre is decreased and the blood pressure falls once more; with the return of medullary anæmia the breathing fails again. It will be noted that the respiratory centre succumbs rapidly to acute anæmia (p. 404) while the vasomotor centre, on the other hand, is for a long while stimulated to increased activity.

This type of periodicity of blood pressure and breathing has been observed in clinical cases of raised intracranial pressure (*e.g.* cerebral tumour, hæmorrhage or abscess).[1]

(2) Periodic breathing may occur in *uræmia* (p. 77) and *heart failure*. In the former condition the respiratory centre is depressed by retained metabolic products, or disturbed acid-base balance. Owing to the feeble state of the centre, shallow breathing results, which in its turn, because of the uneven ventilation of the lungs, produces *anoxia* (p. 447), which gives rise to periodic breathing. First, there is an increase in the breathing which somewhat improves the oxygenation of the centre and washes out CO_2 excessively from the blood. This is followed in turn by dwindling of the breathing and

[1] Eyster, *J. exp. Med.*, 1908, *8*, 565.

apnœa. During this period CO_2 reaccumulates and oxygen lack develops once more and the breathing resumes.

The administration of oxygen to improve the nutrition of the centre, or of CO_2 to act as a respiratory stimulant, abolishes the periods and establishes regular breathing. Improvement of the cerebral circulation resulting from treating the cardiac failure produces more lasting benefit; similarly, renal failure should be suitably treated.

REGULATION OF BODY TEMPERATURE[1]

The normal body temperature at rest is about 97–98° F. (36–37° C.). In health it is always kept fairly close to this level by maintaining a balance between heat gain and heat loss.

Heat Gain.—Heat gain is due to (i) heat *produced* in the body and (ii) heat taken up under certain circumstances from the *environment.*

(1) Heat Production.[2]—As explained on p. 378, heat production under standard resting (basal) conditions is one Calorie per kg. of body weight per hour, or 37–40 Calories (depending on sex) per square metre per hour. This output works out at about 1700 Calories per day in an average man and 1500 Calories in an average woman. Moderate physical activity increases heat production to a total of 2500–3000 Calories per day; if very heavy work is done the total heat output may rise to 7000 Calories or more, per day. Short bursts of extremely severe exercise may increase heat production temporarily to 10–16 times the basal level. The "specific heat" of water is 1, *i.e.* one (small) calorie raises the temperature of 1 g. of water by 1° C. The specific heat of physiological saline, and therefore of the body (which is 80% water) is also, approximately, 1. Therefore, if there were no heat loss, the temperature of the body under basal conditions would rise by 1° C. per hour, and under conditions of normal activity it would rise by 2° C. per hour. But so efficient are the mechanisms for bringing about heat loss that only when work is heavy or when the environmental conditions interfere with the heat loss mechanisms does the body temperature rise well above the normal range.

(2) Heat Gained from the Environment.—The body can take in heat from objects hotter than itself: (i) by direct radiation from the sun or heated ground, or (ii) by reflected radiation from the sky. This type of heat intake is independent of the temperature of the air. The amount of heat gained by radiation can be reduced by wearing garments which reflect the radiations or by making use of any available shade. In the hot desert the body takes up more heat when naked than when covered by thin white clothes.

When, however, the air temperature exceeds that of the skin, the body surface takes up in addition heat from its immediate surroundings; this last sort of heat gain is a great burden to people living in hot climates.

[1] Dill, *Life, Heat and Altitude*, Cambridge, Mass., 1938. Adolph, *Physiology of Man in the Desert*, N.Y., 1947. Malchi and Harch, *Physiol. Rev.*, 1947, *27*, 200. Newton, *Recent Advances in Physiology*, 7th edn., London, 1949, p. 1. Newburgh (editor), *Physiology of Heat Regulation and Science of Clothing*, Phila., 1949.

[2] Cf. footnote 1, p. 375.

Heat Loss.—Heat is lost from the body in several ways:

(i) By *radiation* from the body to cooler objects at a distance.

(ii) By *conduction* and *convection* to the surrounding atmosphere if its temperature is lower than that of the body (or rather, that of the skin). The air in immediate contact with the skin is warmed; the heated molecules move away and cooler ones come in to take their place; these in turn are warmed and so the process goes on. These air movements constitute convection currents.

(iii) By *evaporation of water*. The essential fact to remember is that when 1 g. of water is converted into water vapour, 0·58 Calorie of heat is needed and has to be taken up from the environment; this heat is known as the *latent heat of vaporization* and is measured in Calories per gram. Thus when 1 kg. (approximately 1 litre) of water evaporates, 580 Calories are taken from the immediate surroundings. Evaporation of water takes place from the lungs, and from the skin. *Evaporation of sweat* is the principal means of heat loss when the body temperature tends to rise.

The different methods and routes of heat loss are considered more fully below.

(1) Radiation.—The magnitude of heat loss (or gain) by radiation depends on the *size of the body surface* and on the average *temperature difference* between the skin and the surrounding objects. The part played by radiation, as already indicated, varies widely with climatic conditions. In a temperate climate, a resting person (wearing ordinary clothes) loses about 60% of his heat production by radiation; if he is working, although heat production is increased, there is no significant increase in the absolute amount of heat lost by radiation because the skin and body temperature rises only slightly. (During work the *percentage* of the total heat loss which is due to radiation obviously becomes smaller.)

(2) Lungs.—The expired air leaving the lungs is saturated with water vapour at body temperature; this water is derived by vaporization from the moist mucous membranes of the respiratory passages. The amount of water vapour taken up depends on the initial state of the inspired air; when dry it takes up a good deal but when saturated with water vapour it takes up none at all. On an average (for what such a figure is worth, considering the range of atmospheric conditions) the water loss from the lungs is about 300 c.c. per day, equivalent to a heat loss of nearly 200 Calories (*i.e.* $300 \times 0{\cdot}58$ Calories).

Some body heat is also lost via the lungs by raising the temperature of the inspired air to body temperature; therefore an increase in pulmonary ventilation (especially when the air is dry and cool) increases heat loss.

(3) Skin.—Heat loss from the skin occurs by:

(i) *Conduction-convection* to a degree which varies with the temperature gradient between the skin and the surrounding atmosphere. Skin temperature varies directly with its blood flow; the calibre of the skin vessels, especially those in the hands, feet, and face which are under intense vasomotor control can be nicely adjusted to varying body needs. External cold produces cutaneous vasoconstriction and consequently reduces skin blood flow and decreases heat loss; external heat has the opposite effect (cf. p. 327; Fig. 198).

(ii) *Evaporation of Water. Insensible Perspiration and Sweat.*—Water is lost from the skin (*a*) by insensible perspiration; (*b*) by sweating.

(*a*) *Insensible Perspiration.*[1]—This consists of the passage of water by diffusion through the epidermis (it is called " insensible " because it cannot be seen or felt) ; the fluid lost is not formed by sweat glands. It amounts to 600–800 c.c. per 24 hours, equivalent to a heat loss by evaporation of about 400 Calories. It is produced over the whole body surface at a fairly *uniform* rate and is largely independent of environmental conditions.

(*b*) *Sweating.*—Because of its outstanding importance in temperature regulation this mode of heat (and water) loss requires detailed discussion.

SWEAT GLANDS.—There are two kinds of sweat glands in man.

(i) *Eccrine*, which are distributed generally over the body surface and secrete a dilute solution containing sodium chloride, urea, and (in exercise) lactate. The NaCl concentration is variable, *i.e.* 0·1–0·37%, and depends on the level of *corticoid activity*. Eccrine glands are densest on the palms and soles, next dense on the head, and much less dense on the trunk and extremities.

(ii) *Apocrine*—large glands found mainly in the axillæ and round the nipples, and in the female in the labia majora and mons pubis. They form a secretion of variable composition and of characteristic odour.[2]

Secretion of Sweat.[3]—The sweat glands are supplied by cholinergic fibres present in the sympathetic nerves (p. 481). Secretion is produced by direct or reflex stimulation of the centres in the spinal cord, medulla, hypothalamus or cerebral cortex. Sweat secretion is increased in the following conditions :

(i) With rise of external or of body temperature ; this so-called *thermal sweating* is produced in two ways : (*a*) by the rise of body temperature directly affecting the brain centres, probably in the hypothalamus ; and (*b*) reflexly from the stimulated " warm " nerve endings in the skin.

(ii) In emotional states (*mental sweating*) ; this type is limited as a rule to the palms, soles and axillæ, though in extreme cases it may become more generalized. Mental sweating is due to impulses sent out from the higher centres.

(iii) In *exercise* ; both thermal and mental factors play a part.

(iv) Sweating also occurs (*a*) in nausea and vomiting ; (*b*) in " motion "-sickness, for instance on ships or in aircraft ; (*c*) in asphyxia or anoxia, and (*d*) during normal sleep in children and adults.

In cold environments sweat secretion may be negligible.

The *maximum* rate of sweat secretion in one hour may be as high as 1·7 litres ; 8–11 litres may be lost in 5–8 hours. More usually the 24-hour maximal secretion is 12 litres. In the dry, torrid heat of Boulder City, Nevada (maximal shade temperature 34°–38° C., humidity 6–30%), the sweat

[1] Newburgh and Johnston, *Physiol. Rev.*, 1942, *22*, 1.

[2] Yas Kuno's views on the functions of axillary sweat are interesting, even if they may fail to command general assent. The following is an undistorted summary of his writings. The axillary sweat smells most strongly and peculiarly and is " superior in attractive power to the scent of the sexual organs or of any other parts of the body." It has developed especially in man because owing to the erect posture scents from the sexual organs or from any of the lower parts of the body are not usually perceptible. The axilla is advantageously placed, and so its scent can act more effectively. If it be true that axillary scent has an attraction for the other sex then the fact that it is suddenly discharged in emotional states becomes of great significance. The scented vapour would be retained in the axilla and be disseminated occasionally when the arms were moved, and would not be dissipated rapidly as elsewhere. (The phrasing in the above passage is mainly Kuno's.)

[3] Kuno, *Physiology of Human Perspiration*, London, 1935.

at half-minute intervals, it is found that a constant maximum reading is obtained only after three to four minutes in the mouth, and after a longer time, up to five or even seven minutes, in the axilla or groin (Fig. 271). The maximal temperature in the axilla is lower by 1° F. or more, than that in the mouth (or lower still in thin subjects); the maximal temperature in the

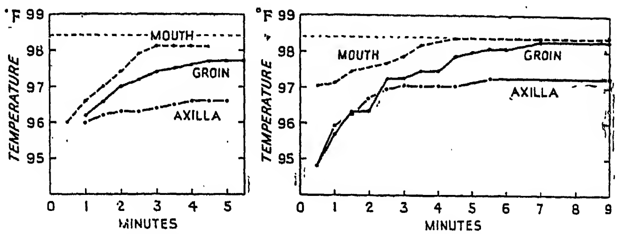


Fig. 271.—Time taken to get a Constant Temperature Reading in Mouth, Groin, and Axilla. (Samson Wright.)

groin approximates more closely to that in the mouth. Rectal temperature is about 0·5–0·7° F., and the temperature in a stream of urine 0·3–0·5° F., above the maximal reading in the mouth.

Effect of Local Conditions on Mouth Temperature—The mouth reading is affected by local conditions:

(i) A misleadingly *high* reading is obtained (*a*) immediately after a hot drink (*e.g.* one or two cups of hot tea (Fig. 272)) when the temperature of the temporarily warmed mucous membrane of the mouth is being recorded (and

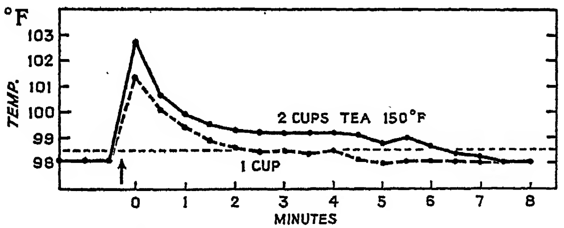


Fig. 272.—Effect of a Hot Drink (one or two cups of tea at 150° F. drunk at point shown by arrow) on Mouth Temperature.
Note the rise of mouth temperature to 101·5–102·6° F.

not that of the blood), or (*b*) after any meal because of the heat produced by the active muscles of mastication and the salivary glands.

(ii) A misleadingly *low* reading is obtained: (*a*) after a cold drink (*e.g.* one litre of ice-cold water (Fig. 273)) which cools the mucosa of the mouth, (*b*) if the nose is blocked, preventing the mouth from being closed; or (*c*) if sweat is evaporating rapidly from the skin of the face and so cooling it.

Diurnal Variation—To obtain useful records of diurnal variation the subject is asked to determine his mouth temperature every hour throughout the day and night for 72 hours, taking the precautions called for from what was said above.

The only way to get regular night readings is to have at the bedside an

alarm clock suitably set to wake the subject up at each successive hour. The value of the temperature readings is considerably enhanced if a careful note is made at the time of what the subject is doing, with special reference to work, food, sleep, room temperature, heart rate, and respiration. Illustrative records are set out in Figs. 274, 275, 276.[1]

Fig. 274 shows that the body temperature may fluctuate during the day between 99·2° and 96·2° F., Fig. 275 between 99·2° and 96·5° F., and Fig. 271 between nearly 98° and 95° F. A diurnal variation of 3° F. may thus occur in any one normal person. The possible extent of the "tolerated" diurnal variation is not always fully appreciated, because it is very unusual to take records at frequent intervals during the night in normal persons,

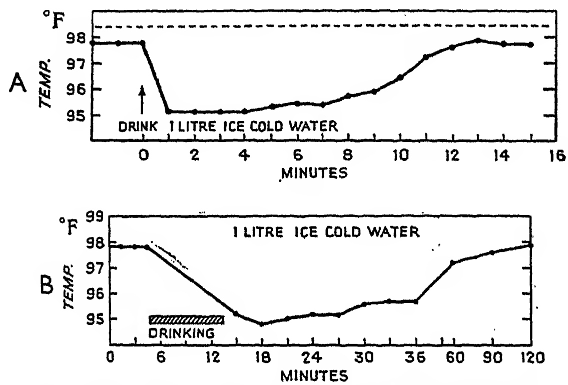


FIG. 273.—Effect of a Cold Drink on Mouth Temperature. (Samson Wright.)

In A the water was drunk as quickly as possible; in B in 12 minutes. The mouth temperature fell to 95° and returned very slowly to normal.

and it is commonly at the inconvenient hours of 2–6 a.m. that the minimum temperatures are observed.

Fig. 276 shows a close correlation between mouth temperature and heart rate, confirming the well-known clinical dictum that the heart rate varies by 10 beats per minute for every 1° F. change in body temperature.[2]

The diurnal variation is fundamentally related to the rhythm of sleeping and waking (Fig. 276). The temperature falls progressively during the earlier hours of sleep, but begins to rise once more in the later hours.

The normal temperature rise during the day is due to muscular activity and associated heat production. During *sleep* the temperature falls, partly because of bodily inactivity but also because of less perfect temperature

[1] Such experiments are of great value in teaching a student the extent of the range of the normal, in himself and his friends, of several so-called "physiological constants." The student also learns to appreciate the fundamental difference between an *average* value and a *normal range*, a distinction which is of the utmost importance in many branches of physiology and medicine.

[2] The curves also illustrate the wide range of the heart rate in a normal person even at rest; the range at rest in this instance is from a so-called "bradycardia" (pulse rate of 60) to a "slight tachycardia" (pulse rate of 90), according to the time of day and the associated body temperature.

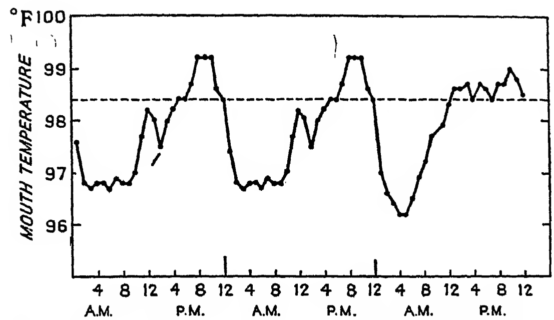


FIG. 274.—Diurnal Variation of Body Temperature. (Samson Wright.)

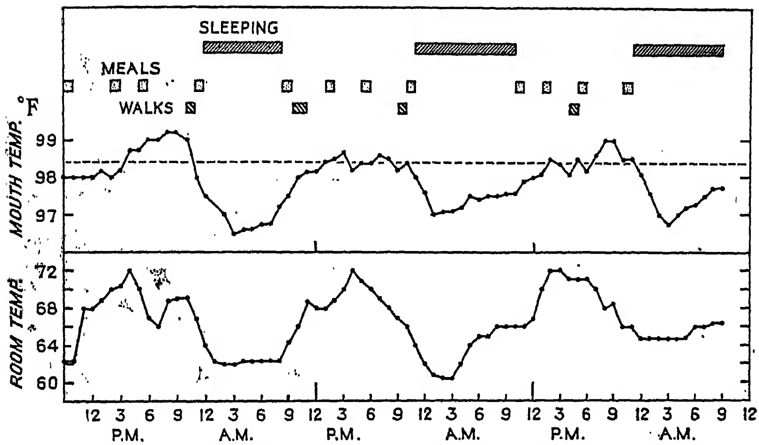


FIG. 275.—Diurnal Variation of Body Temperature in Relation to External Temperature and Bodily Activities. (Samson Wright.)

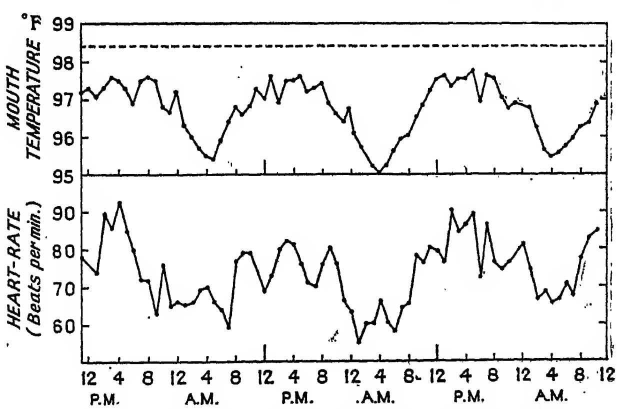


FIG. 276.—Diurnal Variation of Body Temperature and Heart Rate. (Samson Wright.)

regulation. When a healthy man is kept in bed the normal type of diurnal temperature variation takes place, but the evening maximum is not so high.

A reversal of the ordinary daily routine, *e.g.* on going on to night work, ultimately reverses the temperature curve. This does not occur during the first night, for at the hour of the usual minimum (*e.g.* 3–4 a.m.) there is an imperative desire to sleep. After a few days, acclimatization takes place and the sleepiness passes off.

In the tropics the body temperature is about 1° F. above the normal range found in temperate zones. In nervous people, and especially in children, elevations of temperature occur for which there is no adequate explanation.

AGE.—The temperature in *infants* is irregular at first, but periodicity gradually sets in with the development of regular periods of activity and rest. Temperature regulation is imperfect: a fit of screaming causes a rise; a cold bath may lower the temperature by 7° F. In the *aged* the temperature is subnormal: the body is less active, the circulation is feeble, and there is less power of compensation for changes in external temperature. Old people are intolerant of extremes of external temperature.

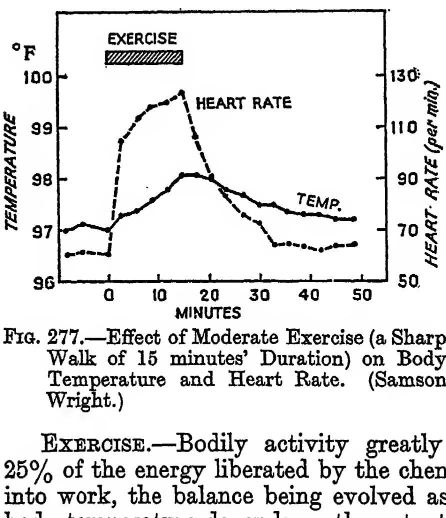


FIG. 277.—Effect of Moderate Exercise (a Sharp Walk of 15 minutes' Duration) on Body Temperature and Heart Rate. (Samson Wright.)

MENSTRUATION.—During menstruation the average temperature is at a minimum. It rises slightly during the next 14 days; the waking temperature (recorded in bed) is said to show a distinct rise at about the time of ovulation.

EXERCISE.—Bodily activity greatly increases heat production; only 25% of the energy liberated by the chemical changes in muscle is converted into work, the balance being evolved as heat (cf. Fig. 265). The effect on body temperature depends on the extent to which heat loss can be increased to make up for heat production; this in turn depends greatly on environmental conditions. After a three-mile race temperatures as high as 103°–105° F. have been recorded in athletes. A sharp 15 minutes' walk may raise the temperature by 1° F. (Fig. 277). When four and six miles are covered in one hour's walking the temperature may rise by 2° and 4° F. respectively. The transient pyrexia is unaccompanied by any ill-effects or feeling of discomfort and is in some ways helpful. It increases the breathing and pulse rate, and facilitates the dissociation of oxyhæmoglobin in the tissues.[1]

[1] In patients with *pulmonary tuberculosis* exercise produces an excessive rise of temperature, which is due in part to the work done, and in part to toxic absorption; this fact is of value in diagnosis. An afebrile phthisical patient and a normal control were allowed to walk up a slight incline of 3¾ miles in 1¼ hour. The average rise of temperature (rectal) in the control was 1·0° F., and in the patient 1·6° F.

RÔLE OF NERVOUS SYSTEM IN TEMPERATURE REGULATION.[1]—Temperature regulation is a highly co-ordinated function depending fundamentally on the activity of the central nervous system. Two mechanisms are involved: (i) reflex; (ii) central.

(i) Stimulation of cutaneous temperature nerve endings reflexly sets up appropriate bodily responses, *e.g.* sweating, vasomotor changes.

(ii) The temperature of the blood bathing the *hypothalamus* directly affects it and likewise sets up appropriate reactions; thus warming the blood in the carotid artery and, more specifically, heating the region of the anterior hypothalamus, results in the manifestations of heat loss, *e.g.* cutaneous vasodilatation and sweating (cf. p. 474). An injury to the *anterior* hypothalamus in animals does not impair the normal reaction to external cold, but annuls the normal increase in heat loss which occurs in hot environments. An injury to the *posterior* hypothalamus likewise abolishes the responses to external heat because the descending fibres from the anterior hypothalamus are interrupted; but in addition the response to cold is also abolished. The anterior hypothalamus is thus the centre for responses to rising temperatures and the posterior hypothalamus the centre for responses to falling temperatures.

The hypothalamus, because of its close connections with the thalamus, doubtless receives all the appropriate afferent nerve impulses; on the efferent side it has access to both the somatic and the autonomic nervous systems and can thus modify muscular and glandular activity, cutaneous circulation, sweat secretion and pulmonary ventilation.

After removal of the *cerebrum* in a dog, normal temperature regulation is preserved but is limited in range. The animal responds appropriately to external heat and cold, but cools down if left for a long time in a cold room. [In clinical lesions of the cerebral cortex changes in skin temperature occur, cf. p. 671.] A *midbrain* preparation (*i.e.* with hypothalamus cut off) is cold-blooded (poikilothermic), *i.e.* it passively follows the temperature of its surroundings.[2]

Whenever the central nervous system is depressed, as by the action of anæsthetics, or grossly injured in essential regions by lesions like hæmorrhage, or if integral parts of the heat-regulating mechanisms are paralysed, *e.g.* the skeletal muscles by curare, the complex and delicately co-ordinated temperature-regulating function fails, and alteration of body temperature results.

[1] Ranson, *Res. Publ. Ass. nerv. ment. Dis.*, 1940, *20*, 342.

[2] As a result of injuries in man to the *pons* and *medulla*, a rise or fall of body temperature may result. The greater part of the body is paralysed and is cut off from the influence of the heat-regulating centres, and therefore follows passively the temperature of the surroundings. If the weather is warm, the room overheated, and the patient well wrapped up and surrounded with hot bottles, the temperature of the paralysed regions rises. No afferent impulses can pass up from this region to the centres at the base of the brain, nor can efferent impulses reach the affected muscles or blood vessels. The temperature of the whole body therefore rises. Similarly, if the paralysed regions are cooled, the general temperature falls.

The effects of *sympathectomy* on the temperature of the limb are described on p. 360. If the *main nerves* to the limb are cut, the temperature rises at first, owing to simultaneous section of the sympathetic constrictor fibres. Later on, the paralysed limb is colder owing to the decreased blood flow which results from the impairment of venous return owing to paralysis of the skeletal muscles; the blood therefore tends to stagnate in the part.

Rôle of Ductless Glands.—On cooling a dog, the metabolic rate increases by 7% before any indication of increased muscle activity sets in; this is probably a measure of the extent to which the ductless glands help in the "fine regulation" of body temperature.

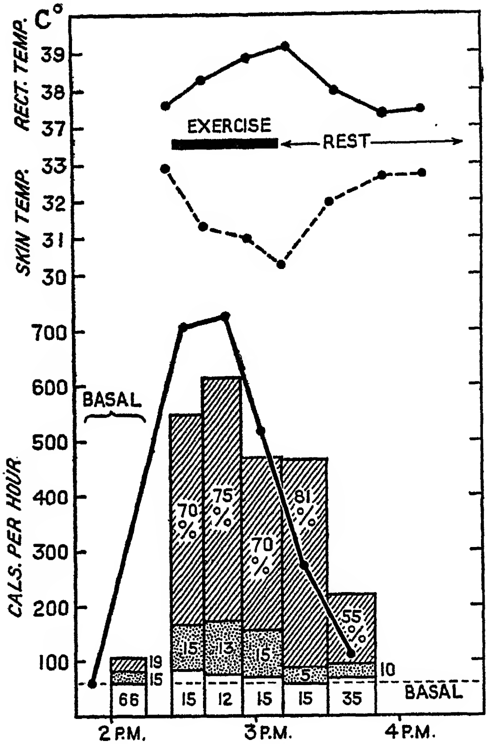


Fig. 278.—Effect of Violent Exercise on Heat Production and Heat Loss. (Du Bois, *Ann. int. Med.*, 1938, *12*.)

Records are: rectal temperature, skin temperature, heat production in Calories per hour (continuous thick line) and heat loss by various channels.
During the period indicated by the thick black line, violent exercise (squash racquets) was carried out for 36 minutes. Heat loss is indicated by columns divided from below upwards into radiation (clear), convection (stippled) and vaporization (oblique lines).
Note the rise of rectal temperature, fall of skin temperature (due to sweating), and increased heat production (from 60 to over 700 Calories per hour). The increased heat loss is due almost entirely to additional loss by *vaporization* (cf. Fig. 265).

(1) *Adrenal Medulla.*—Exposure to external cold reflexly stimulates secretion of adrenaline which stimulates metabolism and decreases heat loss (p. 733).

(2) *Adrenal Cortex.*—See pp. 947 *et seq.*

(3) *Thyroid.*—External cold via pituitary thyrotrophin) stimulates thyroid secretion which increases heat production, mobilizes liver glycogen, and stimulates neoglucogenesis (cf. p. 982).

(i) When an animal (rat) is transferred to a cold environment there are histological changes in the thyroid indicative of functional activity.[1]

(ii) Thyroidectomized animals show impaired temperature control. Patients with myxœdema tend to have a subnormal temperature; the febrile response to the injection of bacterial vaccines is reduced.

Response to Heat.—(1) Effect of Exercise.—When more heat is produced as a result of physical exertion, there is a compensatory increase in heat loss.

(i) The blood flow through the *skin* is greatly increased, leading to a rise of skin temperature, and therefore to a greater temperature gradient between the body surface and the environment. Heat loss

[1] The gland becomes intensely congested; the amount of colloid decreases and loses its affinity for hæmatoxylin; the lining cells become columnar and the mitochondria enlarge and become more distinct. On exposure to an external temperature of 37° C. the gland passes into a "resting" state: the alveoli of the gland are distended with deeply staining colloid, and the mitochondria become indistinct.

by radiation only increases by 15 Calories per hour for each 1° C. rise of skin temperature. Total heat loss by conduction-convection and radiation, though greater than at rest (*e.g.* rising from 75 to 150 Calories per hour) now constitutes only 20–30%, instead of 80% of the total heat loss from the body.

(ii) The *main* heat loss in exercise is due to increased secretion and vaporization of *sweat.* These facts are well brought out in Fig. 278. In this experiment, maximal heat production was at the rate of 700 Calories per hour; heat loss also greatly increased but never exceeded 600 Calories per hour; body temperature, therefore, rose temporarily (cf. also Fig. 265). Sweat secretion in exercise is partly "thermal" and partly "mental" in origin (cf. p. 467); it subsides slowly after the exercise is over.

(2) Effect of a Hot Bath.—The subject gets into a hot bath and

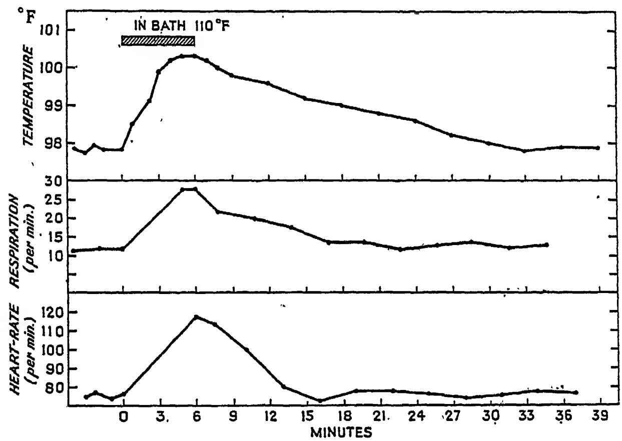


Fig. 279.—Effect of a Hot Bath on Body Temperature, Respiration, and Heart Rate. (Samson Wright.)

stays in it for about five or six minutes; observations are made while in the bath and after getting out of the bath until the control levels are regained. Fig. 279 shows that in a bath at 110° F. the mouth temperature may rise from about 97·8° F. to over 100·3° F.; the heart rate may increase from 75 to 115 and the respiration rate from 12 to 28 per minute. Mouth temperature may subsequently take some 30 minutes to return to normal, though respiration and heart rate recover rather more rapidly. This experiment shows clearly the limitations of normal temperature regulation, and how it can break down quickly under conditions of stress. In a bath at 110° F. heat is being *taken up* by the *immersed* parts of the body, and heat loss depends mainly on evaporation of sweat from the *exposed* parts of the body. As the atmosphere in the bathroom is very humid, evaporation of sweat is probably minimal. Body temperature is therefore raised by the heat released by metabolism as well as by the heat taken up from the surrounding medium. This experiment is of clinical interest in showing that the body temperature may be raised for a considerable time after a hot bath.

(3) Effects of Raised Atmospheric Temperature.—The main automatic compensatory reactions consist (as in exercise) of cutaneous vasodilatation and sweating; adrenaline and thyroxine secretion are inhibited; adrenal corticoid secretion is increased. If the external temperature is lower than body temperature, the increased blood flow through the skin leads to increased heat loss by conduction, convection, and radiation. If the external temperature exceeds body temperature, heat loss can only result from evaporation of sweat. Some of the changes are considered in detail below.

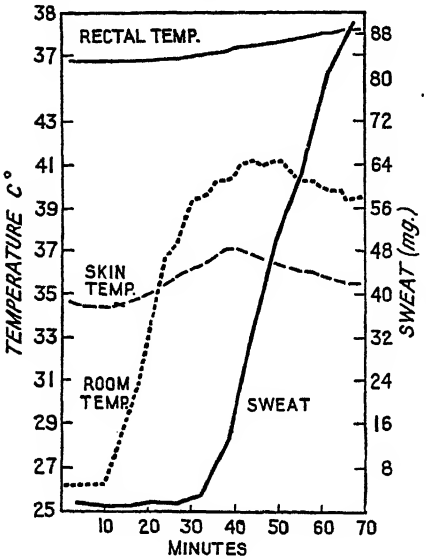


Fig. 280.—Sweat Secretion in Response to Raised External Temperature. (Kuno, *Physiology of Human Perspiration*, London, 1935.)

Ordinate, left: temperature in °C of rectum, skin, and room; ordinate, right: sweat secretion in mg.

(i) *Sweating*.—When the room temperature is raised slowly, *e.g.* from 26° to 41° C., there is a long latent period (*e.g.* 30 minutes) before sweating sets in. Sweating usually commences quite suddenly and coincides with and is due to a rise of *internal* (rectal) temperature (acting centrally) and some rise of *skin* temperature (of the order of 1°–1·5° C. (acting reflexly)). Sweating progressively increases in intensity and may continue to increase even when the room temperature declines; the skin temperature may later fall owing to loss of heat produced by vaporization of sweat (Fig. 280). In hot weather sweating may be produced immediately on exposure to greater external heat.

The importance of sweating in the response to heat is well illustrated by the findings in a patient with *congenital absence of sweat glands*. In the winter his temperature regulation was normal; in the summer, however, it broke down. Thus in July and August his morning temperature was 97° F. and his evening temperature 102·6 F.; pulse rate and respiration followed the temperature changes. The patient and a normal subject were exposed naked in a hot moist room for 30 minutes with the following results:

	Skin Temperature.	Oral Temperature.	Weight loss Skin and Lungs.	Urine Volume.
Normal atmosphere: Patient .	93·0° F.	98·4° F.	..	..
Control .	90·9°	98·4°	..	..
Hot moist atmosphere: Patient	104°	101·4°	22 g.	270 c.c.
Control	99·6°	98·6°	262	10

The patient's skin remained dry and velvety; he felt sick and ill and "panted like a dog"; there was considerable diuresis; pyrexia developed (the mouth temperature rose to 101·4° F.). The control sweated profusely, passed very little urine, and maintained his normal body temperature.

(ii) *Changes in Blood and Circulation.*—The large cutaneous blood flow leads to a decrease in diastolic pressure, an increase in venous return and an increase in cardiac output (p. 282). There is initially an increase in plasma volume. The effect on systolic pressure is variable depending on the extent to which the increased cardiac output compensates for the decreased peripheral resistance. The *pulse rate* always increases even if there is no significant pyrexia (cf. p. 282).

The circulation under these conditions is adversely affected by *standing*. Normally, on passing from the horizontal to the erect position the blood vessels in the legs are reflexly constricted to prevent accumulation of blood in the dependent parts, and to maintain an adequate blood supply to the brain; but during heat exposure vascular tone is dominated by temperature requirements and the normal postural adjustments do not occur. Thus, in hot environments, the subject (*who is strapped to a board*) soon *faints* if he is passively tilted from the horizontal towards the erect position and kept there for some time; in addition his pulse rate further increases (Fig. 282).

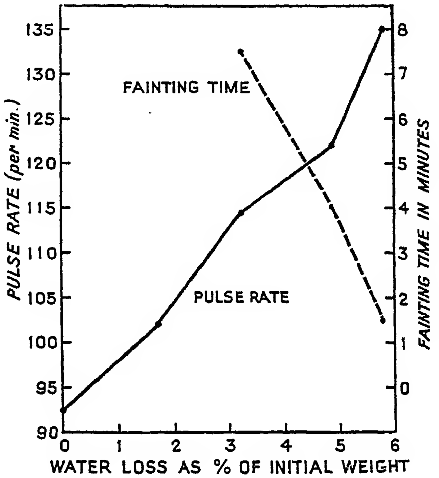


FIG. 281.—Relation between the Degree of Dehydration, the Pulse Rate in the Supine Position and the Fainting Time, when the Subject was passively tilted. (E. A. Adolph, *Physiology of Man in the Desert*, New York, 1947.)

The subject was exposed to a high external temperature and sweated profusely. Note the increase in pulse rate (determined in *supine* position) with increasing dehydration.

Fainting Time: the subject, strapped to a tilt board, was held feet down in the vertical position until he lost consciousness; the time taken is noted.

(iii) *Complicating Effects of Anhydræmia and Salt Loss.*—

(*a*) There is further acceleration of the heart in the supine position (*e.g.* from 92 to 135 (Fig. 281)).

(*b*) The effects of posture are aggravated. Thus in Fig. 282, when 4% of the body weight had been lost (because of water loss) the pulse rate in the supine position was 95; on standing up it rose to over 140.

(*c*) Fainting in the vertical position occurs more rapidly, *e.g.* in 1·5 instead of 7·0 minutes (Fig. 281).

(*d*) When the anhydræmia is severe the plasma volume, the cardiac output, and the blood supply to the skin decrease; as the cutaneous blood supply becomes inadequate, sweat secretion diminishes and may cease. With the main avenue of heat loss thus closed, body temperature rises "explosively" and the general condition becomes precarious.

(iv) *Heat Cramps.*—If the Na^+ and Cl' lost in the sweat are not made good, muscular cramps develop (p. 65).

(v) *Hyperpnœa* develops, the alveolar CO_2 falls and there is, consequently, an alkalæmia which is compensated for by the passage of an alkaline urine, and decreased NH_3 formation by the kidney (Fig. 56).

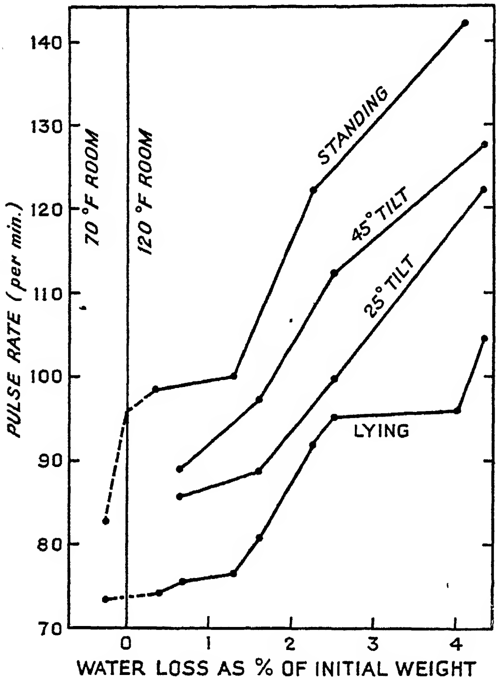


FIG. 282.—Relation between Pulse Rate and Posture in a Man exposed to external heat in whom progressive Dehydration developed as a result of Profuse Sweating. (E. F. Adolph, *Physiology of Man in the Desert*, New York, 1947.)

Control values at room temperature of 70° F.
Experimental values at room temperature of 120° F.
Pulse rates *lying* were taken after 3 minutes in the supine position; *standing* rates, 1 minute after standing up; *tilted* rates, average of readings taken 1, 3, 5, 7, and 9 minutes after being tilted (on a tilt board, feet downwards) at the angle indicated.
Note the progressive increase in pulse rate with dehydration; note also that the pulse rates are higher as the subject passes from the horizontal to the vertical position.

(vi) *Heat Exhaustion.*—This is due to hyperpyrexia, to salt loss, and to dehydration (p. 479).

(vii) *Acclimatization.*—*Prolonged* exposure to external heat leads to "exhaustion" of the sweat glands; the volume of sweat secreted falls and its Na^+ and Cl' content *rises*, *i.e.* the concentration of these ions resembles more closely that of a filtrate from the plasma. On the other hand after *repeated* bouts of exposure to external heat useful adaptations develop: (*a*) sweating commences at a lower level of body temperature; (*b*) the Na^+ and Cl' content of the sweat *falls* with resulting better preservation of the electrolyte content (and crystalloid o.p.) of the extracellular fluids. This latter adaptation is due to increased secretion of the pituitary hormone ACTH which in turn causes a discharge of adrenal corticoids. The latter promote the reabsorption of Na^+ and Cl' from the sweat back into the blood (cf. p. 945).

Rôle of Clothing[1] *and Wet-bulb Thermometer Level in Response to Heat.*—About 20–30 litres of air are entangled between the clothes. In hot weather the clothes retain moisture and the cooling effect of sweating is diminished. Evaporation takes place from the surface of the clothes and only affects the skin through a layer of damp undergarments. Opening the jacket, waistcoat, and shirt gives a feeling of greater comfort, and more moisture can evaporate from the skin surface.

Haldane, attired in flannel drawers, canvas trousers, boots, and stockings,

[1] Buston, *Fed. Proc.*, 1946, *5*, 344.

exposed himself to a hot humid atmosphere (high wet-bulb thermometer temperature). If the wet-bulb temperature was 85° F., the rectal temperature remained normal when the subject was at rest in still air. If the wet-bulb rose above this level, the rectal temperature rose markedly. At 89–90° the rise of body temperature was 1–1·4° F. per hour; at 94°, 2° per hour; at 98°, 4° per hour. In moving air, *e.g.* with an air current of 50 metres per minute, a wet-bulb temperature of 93° F. could be borne without rise of internal temperature. If muscular work was done in still air, the limit of wet-bulb temperature was still lower.

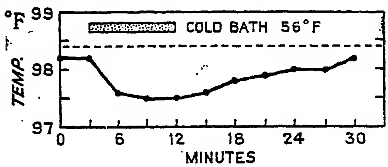


FIG. 283.—Effect on Body Temperature of a Cold Bath at 56° F. (13·3° C.). (Samson Wright.)

HEAT STROKE. — This syndrome develops in men who are working hard, unsuitably clothed, in hot environments. The symptoms are due to hyperpyrexia, salt loss, and dehydration in varying proportions. As the external temperature is high, heat cannot be lost by radiation and convection. If, in addition, the air is moist and still, evaporation cannot take place either. When the heat-regulating mechanism breaks down, the body temperature begins to rise. A vicious circle is thus established with ever-rising body temperature and consequently increasing metabolism and increasing heat production; circulatory failure develops for the reasons already given (p. 477). Individual idiosyncrasy is important; some men do not sweat readily—their skin is hot and burning, but quite dry. Severe and even permanent injury to the central nervous system may occur.[1]

[1] The clinical picture of heat stroke is well illustrated by the following case record, mainly in the doctor's own words: At 7 a.m. I [the doctor] was called to see a man aged 20, who was admitted to hospital the night before complaining of nausea, vomiting, constipation, and lack of energy. He was not regarded as being seriously ill, though his temperature was said to be 101° F.; this was attributed partly to the heat. At 6 a.m. the following day he was found unconscious by an orderly. When I saw him he was in deep coma, responding but slightly to painful stimuli, corneal reflexes absent, pupils moderately dilated and very sluggish to light, breathing somewhat stertorous and rapid, pulse not palpable, not even the carotid. Heart sounds were not heard, skin was cool, rather clammy but no beads of sweat to be seen, and restless purposeless movements. I was told the temperature was 102° F. taken in the axilla. He seemed to me to be about to die. However I took the temperature in the rectum and found it to be 108° F. This simplified the diagnosis and we got to work without delay, first half-covering him with broken ice as he lay in bed, then putting up an intravenous saline. The great saphenous vein when exposed was quite empty and did not bleed even a single drop from either end. Two pints of normal saline were run in. After 45 minutes the rectal temperature was 104·8° F., the radial pulse had returned, restlessness was much less, and the skin of the face, which was growing pink, was now warm, presumably owing to restoration of the peripheral circulation. Within another hour the rectal temperature was 100° F., the pulse diminished in rate and increased in volume. He was able to answer simple questions and with some difficulty to swallow sips of water. After a partial relapse, the patient dramatically improved and by next morning was comparatively normal.

Pyrexial Treatment.—The treatment of rheumatic and other conditions by means of elevation of the body temperature has been extensively practised until recently, especially in America. The procedures employed were: (i) heating the body by means of penetrating high-frequency currents, combined with insulation to minimize heat loss; (ii) placing the subject in a conditioned air-cabinet containing moist air at a temperature of 120°–160° F. Body temperatures up to *107·6°* F. have been maintained for hours by these methods in *robust* subjects. The circulatory strain is great, because the very low diastolic pressure impairs the coronary blood flow (p. 237). In the erect position the systolic blood pressure may be as low as 60 mm.Hg. A simple way of inducing pyrexia is to inject TAB vaccine.

Response to External Cold.—A cold bath (56° F.) may lead to a fall of mouth temperature, *e.g.* of 1° F. (Fig. 283); if a naked man is exposed to an atmospheric temperature of 60° F. for 20 minutes a variable temperature response occurs; the results depend on the efficiency of the compensatory mechanisms. External cold stimulates the metabolic rate owing to increased secretion of adrenaline (p. 733) and thyroxine (p. 474), with resulting increased heat production; if shivering occurs large amounts of heat are released in the muscles. The cutaneous arterioles are constricted to diminish heat loss, and the blood is driven into the deeper parts of the body. The blood pressure may rise. Adrenal corticoids are secreted.

Cold Exhaustion.—This occurs in people who get lost in the cold; the wanderer becomes worn out and surrenders to an overpowering desire to sleep. During this unconscious state, temperature regulation is disturbed and much heat is lost. The body temperature falls; at 68° F. coma sets in and death results. It has been pointed out (p. 328) that a fall of temperature depresses the dissociation of oxyhæmoglobin and tends to lower oxidation in the tissues.

Effects of Refrigeration.—Patients have been deliberately cooled down under anæsthesia to 80° F. for periods up to 36 hours. The blood pressure was unchanged, the pulse rate was slowed, *e.g.* to 60. Usually full recovery occurred on gradually reheating the subject to normal body temperature.

Fever—This term is used to describe the complex response of the body to infection. One of these manifestations is *pyrexia* or raised temperature. The body is believed to be able to deal more effectively with invading organisms when the temperature is raised, and so the pyrexia may be regarded as a protective mechanism. There is increased heat production, but this alone does not account for the pyrexia, because the changes are of an order of magnitude which could be easily compensated for under normal conditions; but heat loss is *decreased simultaneously* fully accounting for the rise of body temperature. Heat regulation in states of fever is therefore *disordered*; body temperature is adjusted to a higher level than normal (as though the hypothalamic " thermostat " were set at a higher temperature). The normal diurnal variation is still present, and the patient responds in the usual way to applications of heat and cold.

In the *initial* stage a rigor or shivering often occurs. The skin vessels are constricted, minimizing heat loss; a rapid rise of temperature therefore takes place. During the *fastigium*, when the temperature is at its height, the cutaneous vessels are relaxed and the skin is flushed; the sweat glands are usually inactive. The rise of temperature increases the metabolic rate by accelerating the oxidative activities of the body; but the heat loss is relatively inadequate to cope with this increased production. During the *defervescence*, when the temperature falls, marked sweating occurs, and heat loss is now greater than heat production.

An agent has been isolated from bacteria, *pyrexin*, which on injection produces fever; pyrexin may be responsible for the complex derangement of temperature regulation just described.

Some of these phenomena are well demonstrated during a malarial rigor. During the shivering attack there is a sudden marked increase in heat production, *e.g.* from 80 to 230 Calories per hour; heat loss, on the other

hand, is unchanged or decreased because of cutaneous vasoconstriction so that all the extra heat liberated is stored in the body. The body temperature may rise to 105° F. as temperature regulation is in abeyance. After an interval the "thermostatic control" is readjusted to 98° F.; sweating sets in, and with increased vaporization there is additional heat loss, bringing the body temperature down to normal once more.

The general manifestations of fever are due to the pyrexia, dehydration, disturbed electrolyte balance in the blood (*e.g.* Na^+ and Cl' loss, alkalæmia), and the action of toxins.

DISTRIBUTION OF HEAT IN THE BODY.—The circulation transports heat to and from the various parts of the body; the distribution of heat within the body, therefore, depends on the integrity of the circulation. In cases of so-called peripheral circulatory failure (*e.g.* hæmorrhage, shock) there may be greater differences than normally between oral, axillary, and rectal temperatures. A low oral temperature may be due to local factors and need not represent a fall of body temperature; the rectal temperature is then the only reliable guide to the temperature of the blood and internal organs. If a hot water bottle is applied to the cold skin, the poor local blood flow may be inadequate to carry away the added heat and the part may be injured if great care is not taken.

NOTE ON INNERVATION OF SWEAT GLANDS [1]

Experiments on animals have shown that the sympathetic nerves which supply the sweat glands are *cholinergic* in type and *not adrenergic*. In support of this work it is found in *man* that drugs which "block" sympathetic adrenergic nerve endings (*e.g.* tolazolin (priscol)) do not inhibit either thermal or emotional sweating. Atropine on the other hand (which "blocks" the parasympathetic (cholinergic) endings) inhibits both these types of sweating. There are, however, some discordant and unexplained findings: intradermal injection of adrenaline (or nor-adrenaline) in *man* does cause local sweating as well as cutaneous vasoconstriction; sweating also occurs when an adrenal medulla tumour is actively secreting (p. 734).

[1] Chalmers and Keele, *J. Physiol.*, 1951, *114*, 510.

V

THE NERVOUS SYSTEM[1]

STRUCTURE AND FUNCTION OF NERVOUS TISSUE

Structure of Nervous Tissue.[2]—The term *neurone* is used to describe the *nerve cell* (*cell body*, *soma*) and its processes, the *dendrites* and the *axon* (*axis cylinder*, *nerve fibre*). The nutrition of the axon depends on its intact connection with its cell.

Nerve cells vary in size and shape in different parts of the body; their cytoplasm is fibrillated, the so-called neurofibrils extending from the dendrites, through the substance of the cell, into the axon. Embedded in the cytoplasm are angular bodies, the *Nissl granules*, which stain deeply with basic dyes like methylene blue; they consist of nucleoprotein and contain organically-combined iron. They are very scanty in many nerve cells. Their size and number vary with the physiological condition of the cell; fatigue, the action of certain poisons, and *section of the axon*, cause the Nissl granules to disintegrate and disappear (*chromatolysis*). The nucleus is large and shows a well-marked chromatin network and nucleoli. The dendrites branch repeatedly immediately they leave the cell, and also contain Nissl granules. Mitochondria and a Golgi apparatus (Fig. 295) can be demonstrated in the cytoplasm.

The *axon* or *nerve fibre* arises from a part of the cell—the *axon hillock*—which is distinguished by the absence of Nissl granules; it runs as a rule for a considerable distance before branching.

A *medullated* nerve fibre consists of the following structures from within outwards:

(i) A central core of semifluid *axoplasm* containing minute rodlets arranged parallel to the long axis of the fibre and giving it a fibrillar appearance (*neurofibrils*). The axoplasm is not covered by a histologically identified surface membrane, but there is a specialized boundary layer of axoplasm, the *axon membrane*, which separates the intracellular contents from the extracellular fluid and is the *functional* surface membrane. It is probably the site of the action potentials recorded from nerve.

(ii) The white *myelin* (*medullary*) *sheath*, which is made up of concentric sheets of protein interspersed between layers of lipoid material (mainly *lecithin*) radially arranged; it stains black with *osmic acid*. The sheath is interrupted at regular intervals by the *nodes of Ranvier*; the distance between the nodes increases during growth, and in the adult animal the internodal distance is proportional to the diameter of the fibre. The functions

[1] See Fulton, *Physiology of Nervous System*, 3rd edn., *N.Y.*, 1949. Fulton, *Textbook of Physiology*, 16th edn., Phila., 1949. Ranson and Clark, *Anatomy of Nervous System*, 8th edn., Phila., 1947. Sherrington, *Integrative Action of Nervous System*, new edn., Cambridge, 1947. *Selected Writings of Sir Charles Sherrington*, ed. D. Denny-Brown, London, 1939. Brain, *Diseases of Nervous System*, 4th edn., London, 1951.

[2] Greenfield, *J. Neurol. Psychiat.*, London, 1938, *1* N.S.), 306.

of the myelin sheath are not known with certainty. Studies with single isolated mammalian medullated nerve fibres suggest that the myelin sheath is an insulating layer; the action potential which accompanies the nerve impulse can only be picked up from such fibres at the *nodes* of Ranvier, where the myelin sheath is absent.

(iii) The *neurilemma* (*sheath of Schwann*), which is a homogeneous membrane closely adherent to the myelin and bent at the nodes. Under it, midway between the nodes, is nucleated cytoplasm, the *cells of Schwann*.

A *non-medullated fibre* consists of axis cylinder, neurilemma, and cells of Schwann, but has no myelin sheath. All the postganglionic fibres of the autonomic system are non-medullated as are those fibres in the somatic nervous system which are less than 1μ in diameter; larger somatic fibres are medullated. It is worth emphasizing that non-medullated fibres are very numerous in the central nervous system and in the dorsal nerve roots.

The surface membrane of the nerve cell and fibre separates two fluids of different ionic content: inside the neurone, the intracellular fluid has a high K^+ and low Na^+ and Cl' concentration; in the extracellular (interstitial) fluid bathing the neurone the reverse obtains (p. 5). The surface membrane in *resting* nerve is almost impermeable to Na^+, but is readily permeable to K^+ and Cl'; this selective permeability is said to be a factor in determining the ionic concentration differences. The differences in ionic concentrations between the exterior and the interior of the neurone (soma and axon) may in some measure be due to the metabolic activity of the entire neurone and not exclusively to the properties of the membrane (cf. p. 8).

Properties of Nerve Fibres.—As a nerve fibre is simply the long process of a neurone, it is somewhat artificial to consider its properties apart from those of the neurone as a whole. But as it is comparatively easy to study the properties of nerve fibres, and more difficult to investigate nerve cells, it is convenient to deal with the nerve fibre as though it were an independent entity. The only function of a nerve fibre is to *conduct the nerve impulse*. Experimentally the fibre can be stimulated at any accessible point on its course; the nerve impulse set up at the point of stimulation travels just as readily centrally as distally. Under natural conditions the nerve impulse is usually generated in the nerve cell and travels distally along the axon; the main exceptions are the afferent neurones (dorsal nerve roots, afferent cranial nerves); in these the impulse is set up by appropriate stimulation of sensory terminals.

Nerve Impulse.[1]—This is an incompletely understood physico-chemical change which is transmitted by nerve fibres; it involves an alteration in electrical state and chemical changes (O_2 consumption, CO_2 output, heat output). The *impulse* must be carefully distinguished from the *stimulus*, which is the external force (*e.g.* electrical, chemical, mechanical) which sets up the impulse. The chemical changes in nerve fibres are probably, wholly or mainly, concerned with the *recovery* processes which follow activity; the *electrical change* is the most certain indicator of the development and propagation of the nerve impulse and probably represents the essential process involved in transmitting the impulse along the fibre (p. 488).

[1] Erlanger and Gasser, *Electrical Signs of Nervous Activity*, Phila., 1937. Symposium on "Physicochemical Mechanism of Nerve Activity," *Ann. N.Y. Acad. Sci.*, 1946, *47*, 375–602. Newton, *Recent Advances in Physiology*, 7th edn., London, 1949, p. 156.

RESTING POLARIZED MEMBRANE OF NERVE FIBRE.—If a single fibre of large diameter is employed (giant fibre of squid, diameter up to 500 μ) a micro-electrode can be introduced into the *interior* of the fibre; another electrode is placed on the *external surface* of the fibre. The electric potential difference between the outside and the inside of the fibre can then be determined by connecting the electrodes to an amplifier and cathode ray oscillograph. The inside of the fibre is found to be at a lower electric potential—*i.e.* it is *negative* compared with the outside (Fig. 284). The surface membrane of the nerve fibre (*i.e.* the axon membrane) is said to be *polarized*, *i.e.* it maintains a potential difference (p.d.) between its outer and inner surfaces; its magnitude (in these giant fibres) is about 50 mV.

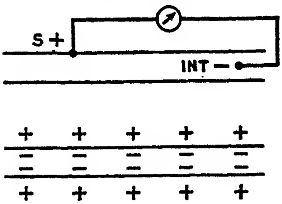


FIG. 284.—Diagram to show Polarization of the Surface Membrane of the Resting Nerve Fibre.

S = surface electrode; INT. = electrode in interior of nerve fibre.
The upper part of the diagram shows how the resting potential has been measured.
INT. is negative relative to S.
The lower part of the diagram shows the distribution of the charge across the nerve membrane. The external surface is positively charged and the internal surface is negatively charged.

The resting potential difference across the axon membrane is attributed principally to the difference of K^+ concentration between the axoplasm and interstitial fluid. To account more completely for the potential difference one must probably also pay attention to the relative permeability of the axon membrane to the three principal ions concerned, viz. K^+, Na^+, and Cl' and to their concentrations, inside and outside the fibre. This is of great importance in connection with the production of the action potential (see p. 485).

INJURY POTENTIAL.—If the interior of the fibre is connected with an *injured* point on the surface of the fibre, the interior and the damaged surface point are found to be at the same potential, *i.e.* they are equally negative. It is supposed that as the membrane has been destroyed at the injured spot, the *surface* electrode now makes connection with the *interior* of the fibre. If one surface electrode is placed on a *normal* point and another on an *injured* point, the latter is found to be negative relative to the former (Fig. 285). The potential difference recorded under these conditions is called the *injury potential*; it is due to connection being made between the exterior and the interior of the fibre, the latter region being, as stated, negative, relative to the former.

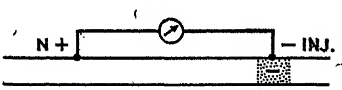


FIG. 285.—Injured Area of Nerve Fibre is Negative relative to Normal Area.

The diagram shows how the injury potential is measured.
N = normal. INJ. = injured areas.
Both electrodes are placed on the external surface of the fibre. INJ. is negative with respect to N.

Effect of Nerve Activity.—A nerve fibre (Fig. 286) is stimulated at S; the impulse set up travels along the fibre towards A; leads are taken from a *surface* electrode at A and another *inside* the fibre (I). When the impulse reaches A, (*i.e.* when "activity" develops at A) it becomes negative relative to I. Activity has changed the state of the membrane at A; the membrane is said to be "depolarized." "Depolarization" means the abolition of the previous state of "polarization"; when the word was introduced, it was thought that at the active spot on the fibre, as at an injured spot, contact

was made with the interior of the fibre owing to a "breakdown" of the membrane which no longer maintained a potential difference between its external and internal surfaces. This view is inaccurate, however. An active spot differs in its electrical state from an injured spot on the nerve fibre. An injured spot has the *same* potential as the interior of the fibre; on the other hand the surface of the active spot is *negative* relative to the interior of the fibre. The membrane at the active spot has thus become *polarized in the reverse direction* (Fig. 286, 3).

Hodgkin and Katz[1] have recently suggested that during excitation of a nerve fibre, the axon membrane changes its relative permeabilities to K^+, Na^+, and Cl'. In the resting state the relative permeabilities are taken to be in the ratio 1·0 : **0·04** : 0·45, and these change to

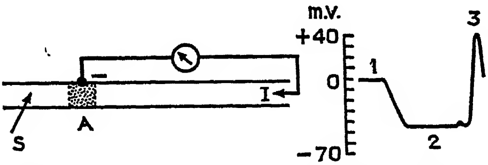


FIG. 286.—Surface of Active Area of Nerve Fibre is Negative relative to Interior of Resting Part of Fibre. Right-hand Part of Diagram shows Electrical Changes:

1 = Isoelectric line (zero potential).
2. Nerve Fibre at rest. Connect electrode on *surface* at A with electrode in *interior* of fibre at I. As I is relatively negative, record moves downwards (−50 mV.).
3. Stimulate at S. As activity develops at A, the surface electrode at A becomes negative compared with the interior electrode, I. Record is deflected above base line (+40 mV.).

Ordinate in millivolts. Upward deflection always represents relative negativity at A (*i.e.* the proximal electrode).

1·0 : **20** : 0·45. Thus the axon membrane temporarily becomes selectively permeable to Na^+ and relatively impermeable to K^+. Na^+ rapidly flows from the interstitial fluid into the nerve fibre until the Na^+ concentration inside the fibre exceeds that outside. This reversal of the resting ratio of external to internal Na^+ concentrations reverses the sign of the resting membrane polarization; as stated above during excitation the surface of the membrane becomes negative relative to the interior.

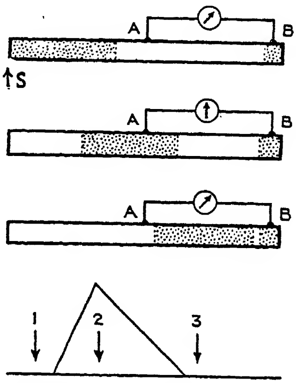


FIG. 287.—Production of Monophasic Spike Potential.

A = Normal area. B = Injured area. The nerve fibre is stimulated at a point S at extreme left. The impulse travels from left to right.
Stippled area = area of activity or injury.
First line: before activity reaches A.
Second line: activity reaches A.
Third line: activity has passed away from A.
Fourth line: resulting electrical record:
1, 2, 3, represent electrical changes corresponding to events in first, second and third line respectively.

With this background we can now consider the normal methods of recording the electrical changes which accompany nervous activity.

Spike Potential (Monophasic Action Potential).—Two electrodes, connected to an oscillograph, are placed on the surface of a nerve fibre at right angles to its long axis (Fig. 287); A is a normal region; B is a region which has been killed or anæsthetized. [The potential difference between A and B (injury potential) can be compensated for.] Stimulate

[1] *J. Physiol.*, 1949, *108*, 37. Keynes, *ibid.*, 1951, *114*, 119. Hodgkin, *Brit. med. Bull.*, 1950, *6*, 322.

the fibre at S; an impulse is generated; when the impulse reaches A the point A becomes momentarily negative compared with its initial resting state. This negativity shows itself as a transient deflection, recorded upwards; (negativity at the proximal electrode is always recorded as an upward deflection). The deflection is a simple upstroke rapidly returning to the base line; it is called, therefore, a *monophasic variation*, or *spike potential* (because of its brevity, shape, and transient nature). The ascending limb (upstroke) represents the development of activity at A; the descending limb (downstroke) represents the dying down of activity at A. The amplitude and the duration of the deflection depend on the diameter of the fibre and

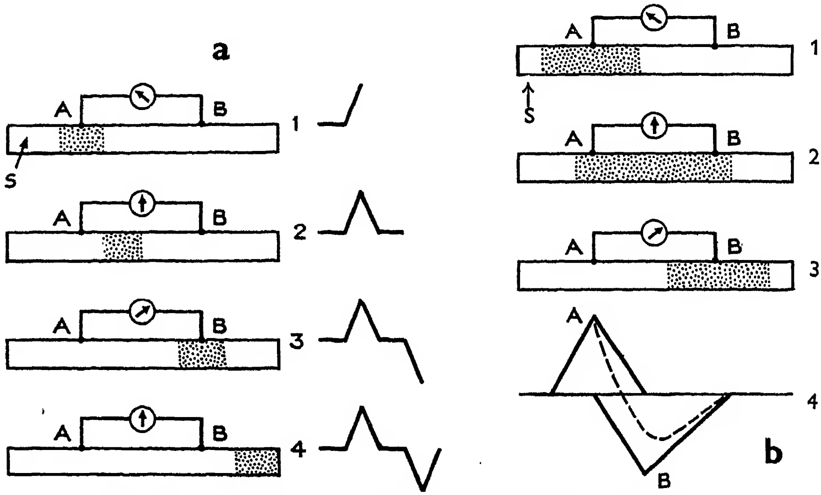


FIG. 288.—Production of Diphasic Action Potential in Nerve Fibre.

Stippled part of nerve fibre=area of activity.
a. Right-hand records: electrical changes. Stimulate fibre at S. 1. Activity develops at A. 2. Activity passes away from A. 3. Activity develops at B. 4. Activity passes away from B.
b. Activity is persisting longer at excited points than in **a**. Stimulate fibre at S.
First line: activity develops at A. Second line: equal activity at A and B. Third line: activity has passed away at A and is still persisting at B. Fourth line: electrical changes:
Activity at A gives rise to a monophasic upward directed spike.
Activity at B gives rise to a similar downward spike. The curve actually obtained is the upstroke of A continued along the interrupted line. It is the resultant of algebraic summation of curves A and B.

other factors. In fibres of large diameter (20μ) the duration of the spike is about 1 millisecond.

During the upstroke of the spike Na^+ leaks into the fibre from the surrounding fluid; during the downstroke K^+ is supposed to leak out of the fibre into the surrounding fluid. During the recovery period the normal resting ionic distribution is re-established.

Diphasic Action Potential.—The experiment is repeated with this difference: the two points A and B from which the leads are taken are both normal. Stimulate at point S (Fig. 288, **a**). The impulse reaches A setting up the usual, upwardly recorded, spike potential which represents the development and subsidence of activity at A. While the spike potential is being recorded at A, the impulse is moving on at its characteristic velocity towards B. Let us suppose that A has already come to rest when the impulse reaches B. B

becomes active compared to A and the spike potential produced here is directed downwards (it is due to negativity at the distal electrode). The passage of the impulse past the two electrodes A and B produces two deflections, the initial one (upwards) and the final one (downwards); this *double deflection* is called the *diphasic variation.* The detailed features of the diphasic variation depend on (i) the duration of the period of activity at A; (ii) the time taken for the impulse to travel from A to B; this in turn depends on (*a*) the distance between A and B and (*b*) the velocity of conduction of the

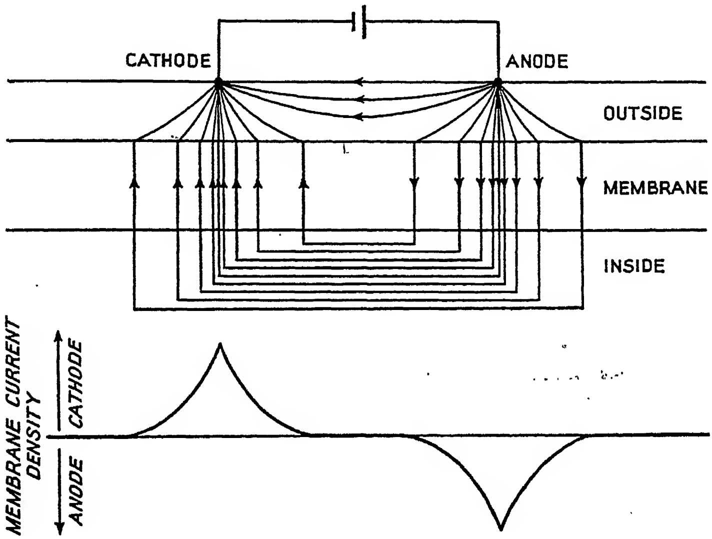


FIG. 289.—Diagram of Current Distribution in Nerve during passage of a Brief Pulse of Current. (After Katz, from Fulton, *Textbook of Physiology*, W. B. Saunders & Co.)

Upper diagram shows the distribution of current flow from anode (+) to cathode (−), outside the membrane (*i.e.* through the interstitial fluid), across the membrane, and inside the fibre (*i.e.* in the axoplasm).

Lower diagram shows membrane current density. It is greatest at the cathode and anode and declines in an exponential manner with distance from the electrodes. Nerve excitability is increased at the cathode (catelectrotonus) and is decreased at the anode (anelectrotonus).

impulse. The diphasic variation in Fig. 288, b, is the algebraic sum of the "overlapping" potential changes (spikes) produced at A and B. When leading from *two normal* points (A, B), the occurrence of a diphasic potential proves that the impulse has been *propagated* from A to B. The occurrence of a *monophasic deflection under these circumstances* proves that the disturbance is *localized* and not conducted. Diphasic variations are sometimes difficult to interpret; monophasic variations led from one normal and one *injured* point are preferable when one is aiming at a detailed analysis of the potential wave and especially when after-effects are being studied (p. 489).

Production of Nervous Impulse.—This is most conveniently examined when the fibre is stimulated by means of a brief pulse of current. Fig. 289 shows diagrammatically how the current flows from the electrode connected

with the anode, through the fluid outside the membrane, through the membrane, to the interior of the fibre, and then out again to the electrode connected with the cathode. The electrical change set up in the neighbourhood of the electrodes is called *electrotonus* : *catelectrotonus* at the cathode ; *anelectrotonus* at the anode. Electrotonus is a *localized* electrical change ; it is maximal at the electrode and diminishes in magnitude (*decrements*) exponentially with distance from the electrodes ; on switching the current on and off it rapidly builds up and subsides.

The following events are observed on increasing the strength of the current. The electrotonus progressively increases in size (Fig. 290). At the *cathode* it is associated with *increased excitability* of the fibre, at the anode with decreased excitability. When catelectrotonus reaches a certain magnitude, an impulse is generated at the cathode as shown by the appearance of a propagated spike potential (Fig. 290 *a, b*). It seems that the current flow at the cathode alters the state of the membrane ; when this so-called " depolarization " (better, change of polarization) reaches a critical magnitude, an impulse is set up.

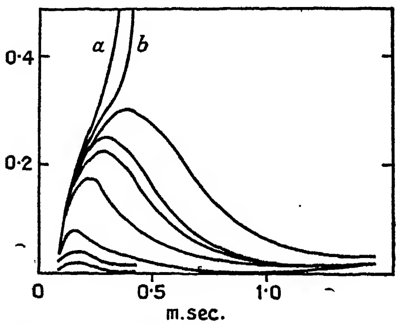


FIG. 290.—Development of Catelectrotonus and Spike Potential in Stimulated Nerve Fibre. (After Hodgkin, *Proc. roy. Soc. B.*, 1938, and Newton, *Recent Advances in Physiology.*)

Electrical changes produced in nerve fibre at the stimulating electrode (cathode) by progressively increased stimuli.

Ordinate : magnitude of response as fraction of full spike potential (*i.e.* fraction of 40 millivolts).

Note the progressively increasing local catelectrotonus. With the strongest stimuli employed (*a, b*) the catelectrotonus develops into a full spike potential of which only the onset is here shown.

To recapitulate : the electrical stimulus first sets up a *localized* electrical change in the fibre (*electrotonus*) ; when the catelectrotonus has risen to a threshold size it generates a *propagated* electrical change (the *spike potential*).

In considering the observations recorded above it should be remembered that under natural conditions nerve fibres are not stimulated midway along their course. In the case of the afferent neurones the nerve endings are stimulated by appropriate stimuli, *e.g.* touch, heat, cold, stretch. It is supposed that the natural stimuli produce the same sequence of changes in the nerve endings, *i.e.* localized catelectrotonus followed by a propagated action potential. In the case of all other neurones the impulse is generated in the cell ; the problem there is, how does an impulse arriving at a synapse stimulate the adjacent cell body ; this question is considered on p. 522.

Propagation of the Impulse.—The *spike potential acts like a travelling, stimulating cathode.* Consider Fig. 291 : at the boundary of the active region of the fibre, *i.e.* at the site of production of the spike potential, a local closed electric circuit is set up, the current flowing from the positive resting region to the negative active region ; this local circuit in front of the active zone is identical with the local circuit set up at the cathode, *i.e.* it is a region of momentary catelectrotonus. When this catelectrotonus rises to the requisite magnitude it alters the membrane sufficiently for an impulse and spike

potential to be set up. The *spike potential acts, then, as the stimulus to the zone of nerve fibre immediately adjacent to it; via the stage of catelectrotonus, the spike potential (and impulse) are progressively propagated along the fibre.*

It follows that the local magnitude of the spike potential depends on the local functional state of the fibre at each point, and not on the magnitude of the exciting stimulus. Wherever a patch of membrane is "depolarized" it produces the maximal spike potential which it can. If a region of a nerve fibre is depressed, *e.g.* by local anæsthetics or injury, the affected region, when stimulated by the arrival of the electrotonic wave (which travels in front of the spike potential) responds with a subnormal spike potential which is the best the depressed region of fibre can produce. If a length of fibre is uniformly but partially narcotized, the spike potential (and impulse) travel through it with a uniformly diminished intensity. When the spike potential reaches normal tissue again, the forerunning electrotonus generates a spike potential of normal magnitude once more.

By means of these circuits an impulse can "jump" a *narrow* region

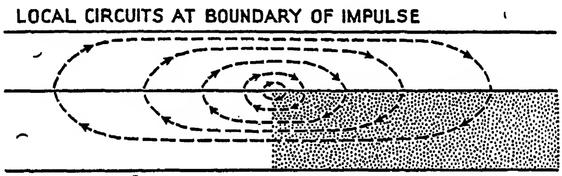


FIG. 291.—Propagation of Nerve Impulse. (After Hodgkin, *J. Physiol.*, 1937, *90*, 183.)

The impulse is moving from right to left in the nerve fibre. Depolarized zone=active area, which is the site of the spike potential. It generates a current flow in the adjacent part of the fibre which sets up first a catelectrotonus and then a spike potential.

of complete block. The spike potential is arrested when it reaches the block; but the local circuits can pass through it to excite the normal nerve fibre beyond. It has been found that if the magnitude of the catelectrotonic potential reaching the normal zone is as small as 10% of the spike potential it can just excite the fibre and produce a propagated spike potential.

AFTER-POTENTIALS.—The spike potential is followed in most fibres by a negative after-potential and then by a positive after-potential. The negative after-potential shows itself by a retardation of the rate of descent of the terminal part of the downstroke of the spike; the positive after-potential consists of a descent below the resting level followed by gradual recovery (Fig. 292). Type B nerve fibres (p. 493) do not display a negative after-potential. The after-potentials are of long duration compared with the spike. During the negative after-potential the excitability of the fibre is increased and the conduction rate is faster; during the positive after-potential the fibre is in a subnormal state, and both excitability and conduction velocity are depressed (cf. Fig. 293).

All-or-None Relationship between Stimulus and Response.—The magnitude of the spike potential (impulse) set up in any *single* nerve fibre is independent of the strength of the exciting stimulus, provided the latter is adequate. An electrical stimulus below threshold strength fails to elicit

a propagated spike potential; if it is of threshold strength or over, a spike (nervous impulse) of maximum magnitude is set up. Either the single fibre does not respond with spike production, or else it responds to the utmost of its ability (under the conditions at the moment). This property of the *single* nerve fibre is termed the *all-or-none relationship*; it should be emphasized that this relationship applies only to the *unit* of the tissue. It applies also to skeletal muscle (the unit being the individual muscle fibre) and to heart (the unit being the entire auricles or the entire ventricles). *Stimuli too weak to produce a spike do, however, set up a local electrotonus*; as already explained, the magnitude of the electrotonic potential progressively increases with the strength of the stimulus until a spike is generated. The all or-none relationship applies only to *spike* production.

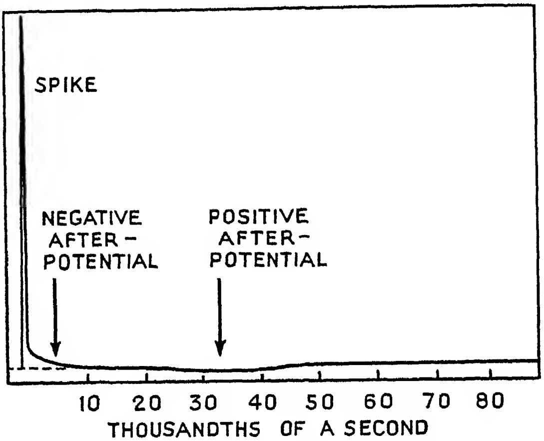


FIG. 292.—After-potentials of Nerve Impulse. (Gasser, *J. appl. Physics*, 1938, *9*, 88.)

Ordinate: magnitude of nerve potentials in arbitrary units.
Abscissa: duration of potentials.

The above account deals, as stated, with the response of a *single* nerve fibre. If a *nerve trunk* is stimulated, then as the exciting stimulus is progressively increased above threshold a *larger number of fibres respond*; the explanation is that the minimal effective (*i.e.* threshold) stimulus is adequate only for fibres of high excitability, but a stronger stimulus is required to activate fibres of lower excitability. Finally a strength of stimulus is reached which excites *all* the nerve fibres; increasing the stimulus further does not increase the response of the whole nerve.

Changes in Nerve during and after Activity.—Changes in excitability and conductivity occur after the appearance of the nervous impulse, and these changes run parallel to one another. (i) For the duration of the spike potential the fibre is *absolutely refractory*. No stimulus, no matter how strong, can initiate a fresh impulse in this region; nor can an impulse generated elsewhere pass through this area. Neither excitability nor conductivity, therefore, is present. (ii) The next few milliseconds represent the period of *partial refractoriness*. At first a very strong current can excite the fibre,

which responds with a subnormal spike. The strength of the minimal exciting stimulus progressively falls until excitability returns to normal.[1] (iii) During the negative after-potential, excitability is supranormal, while during the positive after-potential it is subnormal.

The fibres of largest diameter recover to 90% of the normal in 1 m.sec.; this means that they could conduct 1000 impulses per second of nearly full size, at least for short periods. Motor fibres rarely have to conduct naturally at rates exceeding 100–150 per second; sensory fibres under extreme experimental conditions may have to conduct at 300–400 per second. Normally then these fibres can readily cope with the lower frequencies with which they have to deal.

Fibres of smaller diameter recover more gradually and consequently the upper limit of impulse frequency which they can transmit is lower than in the case of larger fibres.

One of the reasons why a nerve fibre can respond to continuous stimulation for hours without fatigue is that the fibre is conducting not continuously but intermittently; any stimulus falling during the refractory period is ineffective, and the fibre responds again only when it has recovered to some extent.

Metabolism of Nerve Fibres.[2]—Minute quantities of heat are liberated by nerve fibres at rest, during, and after activity.

(i) At *rest*, 2×10^{-5} g. calories are given off per g. of nerve per second; the energy liberated is probably being used to maintain the fibre in its normal excitable state.

(ii) The passage of one impulse through 1 g. of nerve liberates 1×10^{-7} g. calories; this *initial heat* production is very abrupt and coincides with the spike potential.

(iii) During the recovery period there is further liberation of heat (*delayed heat*), which is 8·5 times as great as the initial heat; this energy is used to *restore* the normal excitability of the fibres.

(iv) Increasing the frequency of stimulation raises the total heat production.

(v) In the *absence of oxygen* there is a progressive falling off in the initial heat,[3] a similar falling off in the delayed heat, a progressive but not quite parallel decrease in the size of the spike potential, and after 2–3 hours, extinction of all activity in the nerve. When *oxygen is readmitted* the capacity for heat production returns and recovers fully in 1–2 hours.

(vi) Nerve fibres *use oxygen and liberate CO_2*; the respiratory exchanges are increased by activity.

Metabolism of Neurones.—The heat production of nerve *fibres*, as indicated above, is extremely small. Hill has calculated that stimulated skeletal muscle gives off about 7000 times the amount of heat released by an equal weight of nerve fibres stimulated for the same time. It should be remembered, however, that a nerve fibre is merely the long process of a *neurone*. The metabolic rate of neurones, unlike that of nerve fibres, is extremely high.

[1] Similar changes occur in *heart* muscle though the sequence of events is much slower (cf. p. 235).

[2] Hill, *Chemical Wave Transmission in Nerve*, Cambridge, 1932.

[3] In *muscle*, the initial heat is unaffected by oxygen lack, while the delayed heat is almost completely abolished (cf. p. 431).

Interesting comparisons can be made between brain and skeletal muscle. The blood flow to the brain (p. 306) is 54±12 c.c./100 g./min.; the arterial-venous O_2 difference is 6·5 c.c./100 c.c.; the O_2 consumption of brain is thus 3·5 c.c./100 g./min. The blood flow to violently contracting skeletal muscle (p. 433) is, say, 40 c.c./100 g./min.; let us assume that the venous blood is completely reduced, the arterial venous O_2 difference thus being 19 c.c./100 c.c. The O_2 consumption of muscle is then 7·5 c.c./100 g./min. But the brain contains much white matter, *i.e.* nerve fibres, with a low metabolism and a poor blood supply. Taking this fact into account, it appears that the metabolic rate and blood supply of cerebral *grey* matter is of the same order of magnitude as maximally contracting skeletal muscle, a surprising but important conclusion.

The blood supply of grey matter is not related to the number of nerve cells present, but to the wealth of synaptic connections. Thus a dorsal root ganglion which contains many nerve cells but no synapses has a relatively small blood supply. It would seem, therefore, that the level of metabolic activity at synaptic terminals (presumably in relation to transmission processes) is very high.

The oxygen supply of the brain, though so large, *is not greatly in excess of cerebral needs*; any reduction in the flow or in the oxygen tension or content of the blood soon leads to disturbed brain function. The metabolism of the brain (and of nervous tissue generally) is peculiar. It can apparently only metabolize carbohydrate; but as its glycogen content is very small it is dependent for its energy on its contemporaneous blood sugar supply. A fall of blood sugar (hypoglycæmia) in man produces predominantly symptoms of cerebral dysfunction (p. 915). Lack of the vitamins which participate in the enzyme systems concerned with carbohydrate breakdown (*e.g.* thiamine, p. 1026) produces derangement of the nervous system and ultimately permanent damage.

Different regions of the brain respond in a distinctive manner to physicochemical changes in their environment. Thus the respiratory centre is markedly stimulated by CO_2 excess or a rise of H^+ ion concentration; apomorphine acts specifically on the vomiting centre; cells in the hypothalamus are sensitive to minute variations in blood temperature; morphine seems to act at the thalamic level; in fact much of the pharmacology of the central nervous system is an expression of the individual characteristics of different groups of nerve cells, which may prove to be associated with metabolic peculiarities.

Types of Nerve Fibres.—Three main groups of nerve fibres are recognized:

A fibres: all medullated somatic nerve fibres whatever their diameter (range 1–20 μ).

B fibres: medullated (*i.e.* preganglionic) fibres of the autonomic nervous system (diameter range 1–3 μ).

C fibres: all non-medullated fibres, somatic or autonomic (diameter under 1 μ).

Properties of A Fibres.—These have been studied in greatest detail.

(i) *Conduction velocity*: this is directly proportional to fibre diameter; fibres of 1 μ conduct at the rate of 5 metres per sec.; the largest fibres

of 20 μ diameter conduct at the rate of 120 metres per sec.; fibres with intermediate diameters conduct at intermediate velocities. (Roughly, the diameter in μ, multiplied by 6, equals the velocity in metres per sec.)

(ii) The height of the spike potential is directly proportional to the diameter of the fibre.

(iii) Fibres with the largest diameters are most excitable, *i.e.* they respond to weaker stimuli than thinner fibres.

(iv) The time relationships and characteristics of the phases of the recovery process are shown in Fig. 293. After the initial full recovery there is first a phase of increased excitability corresponding to the negative after-potential, followed by a second phase of decreased excitability corresponding to the positive after-potential.

Properties of B and C Fibres.—B fibres (unlike A or C fibres) develop no negative after-potential and therefore do not show the associated phase of supra-normal excitability. C fibres resemble A fibres in their general characteristics but all the events are on a *smaller scale* and *greatly drawn out in time* (Fig. 293).

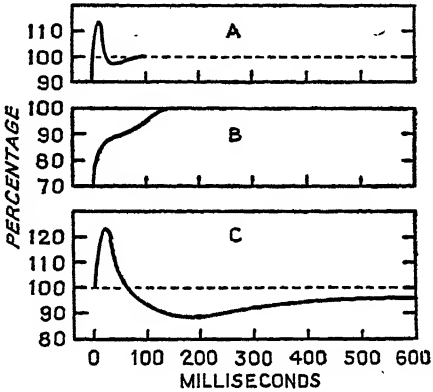


Fig. 293.—Recovery of Excitability of A, B, and C Nerve Fibres, after Single Responses. (Gasser, *Ohio, J. Sci.*, 1941, *41*, 145.)

A, B, C=A, B, C fibres respectively.
Ordinate: Level of excitability. 100=Control level of excitability.
Value over 100=supranormal state; value under 100=subnormal state.
Abscissa: the time interval between the stimulus which produced the response and the subsequent "test" stimuli.

Time Factor in Excitation.—In considering the effectiveness of a stimulus, attention must be paid not only to the strength of the current, but also to the time during which it is allowed to flow through the tissue. A strong current produces a response after a very short time of flow; as the strength of the current is reduced a longer duration of flow is required to stimulate; if the current falls below a minimum value it fails to stimulate no matter how long the duration of flow. The relationship between the current strength and the duration of current flow necessary to stimulate is shown in a *strength–duration curve* (Fig. 294).

Chronaxie and Rheobase.—The weakest current strength which can excite a tissue if allowed to flow through it for an adequate time is called the *rheobase*. (A still weaker current fails to stimulate no matter how long the duration of current flow.) The tissue is then tested with a current strength equal to *twice* the rheobase; the *duration* of the current flow which stimulates the tissue under these conditions is called the chronaxie (Fig. 294).

Chronaxie values are a useful index of the relative excitability of tissues; thus the chronaxie of nerve fibre is considerably shorter than that of skeletal muscle. (The changes in the strength–duration curve and chronaxie of *denervated* skeletal muscle are considered on p. 506.)

ACCOMMODATION.[1]—A current which rises *suddenly* to its full value is more effective and excites at a lower threshold than if it rises slowly. In the case of nerve fibre, the current must have a minimal gradient, in order to excite : if the current rises more slowly than this, it is ineffective no matter what strength it ultimately attains. It is suggested that the tissue, if allowed sufficient time, sets up a " resistance " to the current and so to say " accommodates " itself to it. The current seems to set up some process which counteracts the stimulus and consequently increases the threshold at which excitation can occur. The nature of this process is unknown, but it is called ' accommodation."

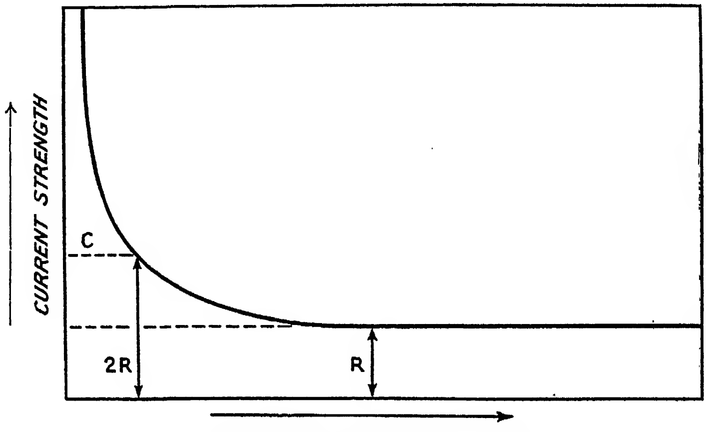


FIG. 294.—Strength-Duration Curve. Rheobase and Chronaxie.

The curve shows the relationship between the *strength* of stimulation (current strength) and the minimum *duration* of stimulation (*i.e.* duration of current flow) necessary to set up an impulse.
R=Rheobase.
2R=2×R.
When 2R is the strength of stimulus, the minimum duration of current flow which can produce excitation is called the chronaxie (C).

To determine the accommodation the tissue is stimulated with currents which rise at different rates ; the threshold strength is noted. If accommodation is rapid (or " good ") the threshold strength with a given rate of current rise is high : if accommodation is slow (or " poor ") the threshold strength with the same rate of current rise is low. If accommodation is lacking the threshold will be at the same level irrespective of the rate of current rise.[2] Impaired accommodation is probably responsible for the phenomena of *tetany* (p. 1004).

Effects of Nerve Section.[3]—**1. Changes in the Nerve Cell.**—*Chromatolysis* occurs in the nerve cells giving rise to the cut fibres. The changes begin within 48 hours of section of the nerve, and reach their zenith in 15 to 20 days. The Nissl granules break up into a fine dust ; later the dust

[1] Kugelberg, *Acta physiol. scand.*, 1944, *8* ,Suppl. 24.
[2] A similar phenomenon in nerve endings is called " adaptation " (p. 550).
[3] Cajal, *Degeneration and Regeneration of Nervous System*, 2 vols., London, 1928. Young, *Physiol. Rev.*, 1942, *22*, 318.

may lose its staining reaction, and the cell appears colourless (Fig. 295). The Golgi apparatus fragments and dwindles in amount. The cell swells from increased fluid content, becomes rounded and the neurofibrils disappear. The nucleus is displaced to the cell margin; it may be completely extruded, in which case the cell atrophies and finally disappears.

The degree of chromatolysis depends on (i) the proximity of the site of section to the nerve cell; (ii) the nature of the section. If the nerve is forcibly torn, death of the cell often results.

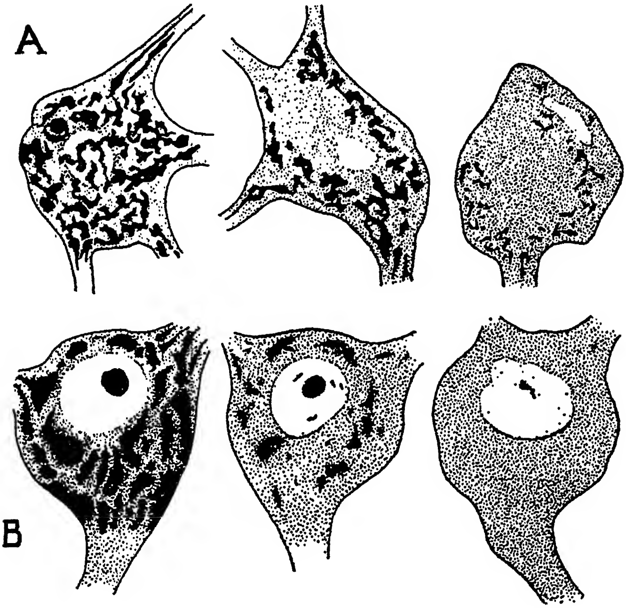


Fig. 295.—Changes in Nerve Cells of Spinal Cord during Chromatolysis. (After Penfield and Dolley, from Fulton, *Physiology of Nervous System.*)

A. Golgi apparatus. From left to right: normal cell; break-up of apparatus 7 days after section of axon; dissolution of apparatus after section of spinal cord.
B. Nissl granules. From left to right: normal cell; moderate chromatolysis from fatigue; extreme chromatolysis.

Repair begins in 20 days and is completed in 80 days. The Nissl granules and Golgi apparatus gradually reappear; the cell regains its normal size, and the nucleus returns to its central position. Cell repair may occur even if the axon does not regenerate.

Afferent Neurones.—(i) If the *peripheral* axon of a dorsal root ganglion is divided, the changes described above occur, but pass away more rapidly. (ii) If the *central* axon is cut, the changes in the ganglion cells are slight. [Similarly, in tabes, where the central axons of the dorsal nerve roots are damaged by syphilis, the ganglia show very slight changes.] The *synaptic terminals* in the spinal cord (p. 522) show characteristic changes; they swell after 24 hours, begin to disintegrate after 3 days, and disappear in 6 days.

Central Nervous System.—Most of the affected cells atrophy completely. As usual the atrophy is more intense when the fibres are cut close to their parent cells.

Changes in *autonomic* ganglia are difficult to demonstrate owing to the normal sparseness of the Nissl granules.

Chromatolysis also occurs in *exhaustion, intoxications,* and *decreased blood supply.* In *hyperpyrexia, e.g.* brief exposure to 109° F., or more prolonged exposure to 107° or 108° F., death of the cell and coagulation of its substance result.

2. **Changes in the Nerve Fibre.**—(1) Stage of Degeneration.—The part of the nerve fibre *distal* to the point of section undergoes degeneration. The process sets in within 24 hours and on the third day the fibre ceases to conduct impulses. Histological and chemical changes occur simultaneously along the whole length of the distal stump. The axis cylinder splits up into short lengths : the medullary sheath breaks up into oily droplets of varying size. During this period the fibre stains black with Marchi's fluid (p. 499). Degeneration may occur in the central end for a variable (usually short) distance from the point of section. Degeneration is complete in 3 weeks ; the fibre debris is then removed by surrounding macrophages which penetrate into the neurilemmal tubes (Fig. 296), the process being complete in 3 months.

Soon after the section, the nuclei of the Schwann cells divide mitotically along the whole course of the *peripheral* stump ; the responsible stimulus is unknown. The Schwann cytoplasm likewise increases in amount and gradually fills the neurilemmal tube as the debris is removed. The Schwann tissue differentiates into thin elongated cells which grow from the cut end of the distal stump and sprout in all directions, progressing at the rate of 1–2 mm. per day. The "useful" growth is towards the central end of the cut nerve ; in this way the Schwann tissue may bridge a considerable gap between the two ends of the cut nerve, even up to 3 cm., and establish connections with the central end. If the two cut ends of the nerve are *stitched together* the Schwann cells readily establish a fibrous bond of union. The peripheral tubes shrink to half diameter in 7 weeks and may so persist for about 18 months.

(2) Stage of Regeneration (Figs. 296–298).—At the same time the *central* axons elongate and grow out in all directions. Each growing fibre splits and may give rise to many, even up to 100, fibrils. The regenerating fibrils meet the Schwann cells which serve as a mechanical bridge to guide the fibrils to the peripheral neurilemmal tubes. As far as is known chemical attractive forces are not involved. Two or three weeks after the section, the peripheral tubes contain varying numbers of developing fibrils, some none, some as many as 25, attached to their inner wall and steadily growing distally (Fig. 296). Some fibrils may pass *between* the tubes and somehow receive a covering of Schwann tissue. Finally all the fibres in any distal tube degenerate except the *single successful one* which progressively enlarges to fill the tube. Fibrils from one fibre may enter several tubes but probably only one fibril can survive. The daily rate of growth is 0·25 mm. in the junctional area of scar tissue, 3–4 mm. in the peripheral stump, and rather faster in the distal end of a *crushed* nerve where the mechanical conditions for regeneration are easier than in a cut nerve. Medullary sheaths begin to develop in about 15 days and follow the course of the fibre (Fig. 297). Increase in fibre diameter takes place very slowly

(Fig. 298). The diameter finally attained is limited by the diameter of the peripheral tube and the size of the parent nerve cell. A thick fibre cannot develop in a thin neurilemmal tube; the development of thick fibres in

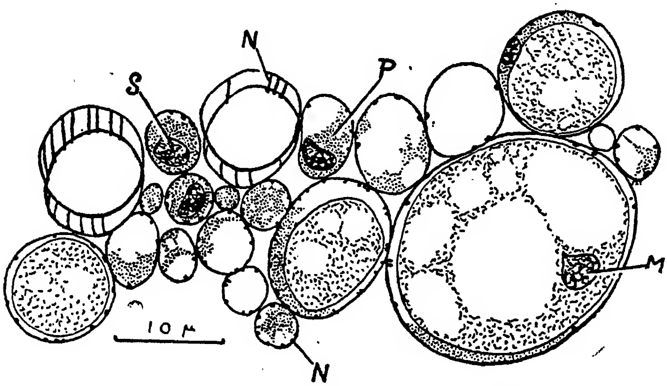


FIG. 296.—Transverse Section of Fibres in Peripheral Stump of Nerve, 2 cm. distal to site of Section. The nerve had been cut 25 days previously and the cut ends immediately sutured. (Young, *Physiol. Rev.*, 1942, *22*, 338.)

N = nerve fibrils, some cut longitudinally and some transversely (latter appear as coarse dots); S, P = nucleus and protoplasm respectively of Schwann cell; M = macrophage nucleus and related cytoplasm. Macrophages have invaded the Schwann tubes and compress Schwann nuclei (*e.g.* right-hand tubes). Many fine nerve fibrils run along the inner aspect of the membrane lining each tube.

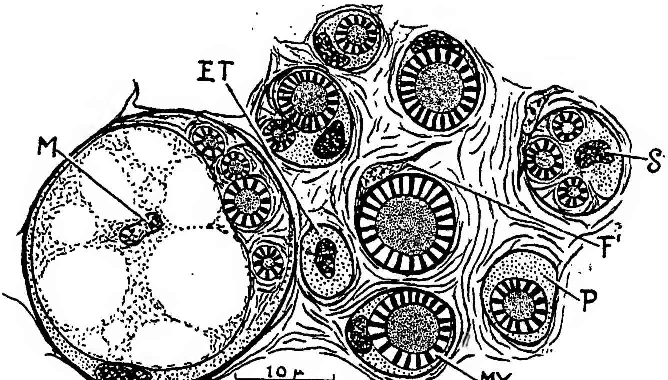


FIG. 297.—Transverse Section of Peripheral Stump of Rabbit's Nerve severed 150 days previously. The stumps were left unsutured and union was made by outgrowth. Medullation is consequently somewhat retarded. (Young, *Physiol. Rev.*, 1942, *22*, 341.)

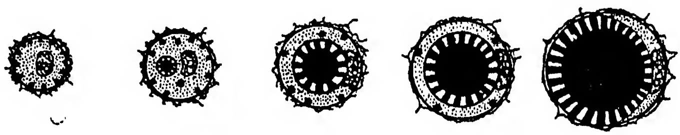
M = macrophage nucleus; P = related cytoplasm; S = Schwann nucleus and related cytoplasm; My = myelin of developing nerve fibres; F[1] = fibroblast. Note that the Schwann tubes contain one to four nerve fibres of varying size with developing myelin sheaths. Only one tube (ET) is uninnervated.

general is hampered by the shrinkage of the peripheral tubes to which reference was made above; a small cell can only give rise to a fine fibre even when the latter finds itself in a large tube. Regenerating rabbit nerve 2 cm. below the

point of suture examined 200 days after the section showed very few fibres exceeding 2 μ in diameter.

The frameworks of the sensory and motor *endings* can persist for months. When the growing axon tip ultimately reaches the nerve ending it establishes a connection with it which at first may be atypical in appearance but later becomes normal. A good many fibres will establish connections with *new kinds of endings in new situations*; the functional complications that may result are discussed on p. 556. It is possible that when a number of nerve fibrils grow down one tube to the same ending the appropriate one alone survives to establish effective connections. It is conceivable also that when a new fibre reaches skin or muscle it may call forth locally the appearance of a suitable ending. [In tissue culture, when a nerve fibre meets a myoblast a motor end plate is formed.] As is explained later (p. 557) functional recovery lags considerably behind anatomical regeneration.

Regeneration never takes place in the central nervous system; regeneration of the central axon of a dorsal nerve root may occur as far as the pia, but there is no penetration into the spinal cord.



25 50 100 150 ?400

FIG. 298.—Stages in Progress of Regeneration in Schwann Tube after immediate Suture in Rabbit. (Figures show time after suture in days.) (Young, *Physiol. Rev.*, 1942, *22*, 343.)

25 days: many fine fibres at edge of tube.
50 days: one or two fibres have enlarged and are placed more centrally.
100–150 days: one large medullated fibre occupies centre of tube, the fine fibres still remaining at the periphery.
400 days: fine fibres have disappeared; central fibre has further enlarged.

Transneuronal Degeneration.—A nerve cell may sometimes undergo degeneration if the *afferent* fibres passing to it are cut. Thus when the optic nerve is cut, the cells in the lateral geniculate body (round which the optic nerve fibres end) undergo chromatolysis; these cells, however, receive afferents from no other neurones except those arising in the retina (p. 575). But similar degeneration may occur after dorsal root section in the cells in the dorsal horn of the spinal cord which receive collaterals not only from the dorsal nerve root but also from descending fibres. The ventral horn cells may degenerate in cerebral tumours involving the precentral convolution although ventral horn cells receive afferents from very many sources. Transneuronal degeneration may be a factor in the pathogenesis of system diseases, *e.g.* when both the pyramidal tract and the lower motor neurones are involved.

Staining Reactions of Nerve Fibres.—(1) OSMIC ACID.—Normal medullated nerve fibres stain black with osmic acid which is reduced to the black osmic oxide.

(2) WEIGERT-PAL METHOD.—For accurate investigation of the white matter of the nervous system, the tissue must be fixed and hardened for a long time (weeks) in Müller's fluid (weak bichromate solution). To

demonstrate areas of sclerosis or regions where the nerve fibres have been destroyed and have completely disappeared, the *Weigert-Pal* stain is employed. The principle of the method is to stain the whole section black with a modified hæmatoxylin stain. The stain is then removed from the diseased (scar) areas by means of bleaching agents like potassium permanganate or sodium sulphite. The *normal fibres stain a deep blue-black,* and the *scar areas remain quite colourless.*

(3) MARCHI'S METHOD.—Marchi's fluid (a mixture of potassium bichromate and osmic acid) does not affect normal nerve fibres, but stains the sheaths of fibres in the *early stages of degeneration* black. The reaction can be obtained between the 8th and 21st day and is maximal about the 12th day. The mode of action of the stain is obscure.

(4) Special stains must be employed to demonstrate *non-medullated* fibres. These techniques have revealed the very widespread distribution of non-medullated fibres in the *central* nervous system.

PROPERTIES OF THE MOTOR UNIT[1]

The Motor Unit.—Skeletal muscles receive their motor nerve supply from the ventral horn cells of the spinal cord and the corresponding cells in the motor cranial nuclei. Each ventral horn cell (or cranial equivalent) supplies a considerable number of muscle fibres, varying with the individual muscle, from 5 to 150. A ventral horn cell and its efferent fibre is a *motor neurone* ; it is called by clinicians the *lower motor neurone.* A motor neurone together with the group of muscle fibres which it innervates is called a *motor unit.* The smallest group of muscle fibres that can ever be employed naturally in the body, either in reflex or voluntary activity, is obviously that supplied by a single motor neurone. The muscle fibres in one motor unit when thrown into tetanic (maximal) contraction yield a tension of 5–30 g. ; this then is the tension range that can result from the maximal discharge of a single ventral horn cell. A single *twitch* of a motor unit yields much smaller tensions (one-fifth the tetanus tension). The muscle fibres in the larger motor units are supplied by thicker fibres (*e.g.* 15 μ), the smaller by finer fibres (*e.g.* 4 μ). The size of the unit varies inversely with the precision of the movements performed by the part ; *e.g.* in the limb muscles the unit may contain 150 muscle fibres, in the eye fewer than 5.

All the efferent fibres passing to skeletal muscle are excitatory, *i.e.* they produce *contraction* of the muscle fibres. There are no efferent nerves which on stimulation produce relaxation or elongation of the muscle, *i.e.* there are no inhibitory efferents. [In the case of smooth and cardiac muscle the efferent supply is both excitatory and inhibitory.] Skeletal muscle contraction under natural conditions always results from discharge of the motor neurones ; muscular relaxation is the result of a decrease or cessation of discharge of the motor neurones (cf. p. 1127).

Response of Muscle to Motor Nerve Stimulation.—(1) A *single* electrical shock of adequate strength applied to the motor nerve gives rise to a

[1] Fulton, *Muscular Contraction and Reflex Control of Movement,* London, 1926. For recent work on muscle action potential see Hodgkin and Nastuk, *J. cell. comp. Physiol.,* 1950, *35,* 39 ; Hodgkin, *Brit. med. Bull.,* 1950, *6,* 322.

simple muscle twitch (Fig. 299). After a short latent period the muscle contracts and then relaxes; the twitch is over in about 0·1 second. If the stimulus applied is of threshold strength it can (theoretically) excite one motor nerve fibre and its related group of muscle fibres, *i.e.* one motor unit. As the strength of the stimulus applied to the nerve is increased, more nerve fibres, and therefore more groups of muscle fibres respond, and the strength of the resulting contraction is correspondingly greater. With what is called *maximal* stimulation, all the nerve fibres are excited and consequently all the muscle fibres supplied by the motor nerve contract.

(2) If a *second* maximal stimulus is applied at varying intervals after the first, the following results are obtained. If it falls during the latent period of the muscle (*i.e.* the first few milliseconds), the arrival of the nervous impulse produces no additional response—the muscle is said to be completely refractory. If applied later, a second muscular response results, leading to further development of tension (*summation of effects*) irrespective of the phase

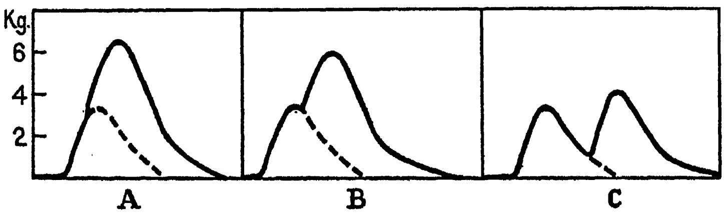


FIG. 299.—Summation of Responses in Skeletal Muscle. (After Cooper and Eccles, *J. Physiol.*, 1930, *69*.)

Isometric Contraction Curves of Mammalian Nerve-Muscle Preparation.
Ordinate = tension in kg.
In each record the lower curve represents the response of the muscle to a single *maximal* stimulus to the motor nerve. The continuous thick line represents the response to the initial stimulus followed by a second stimulus. The second stimulus in A was applied during the rise of tension, in B at the height of contraction, and in C during relaxation. Further development of tension occurred in each case.

of the muscle cycle in which the stimulus is applied, *i.e.* whether during rising tension, the height of contraction or relaxation (Fig. 299). The tension resulting from two maximal stimuli applied successively at suitable intervals may thus be considerably *greater than that from a single stimulus of the same strength.*

(3) If a *series* of maximal stimuli are applied at increasingly short intervals, increasingly complete degrees of summation take place. At low rates (in mammalian muscle, 10–20 stimuli per second) the mechanical fusion is incomplete and the muscle gives a tremulous response (*partial* or *sub-tetanus*). At higher rates (about 60 per second) the mechanical fusion is complete, and *full tetanus* results (Fig. 300). It is essential to appreciate that the more complete the tetanus the *greater is the tension* exerted by the fibres, and the *steadier their pull*; in fact, with full tetanus it may be almost impossible to detect the slightest flicker in the mechanical record, though the electrical record shows a series of discrete waves corresponding to the arrival of each nervous impulse (Fig. 302, B). As *maximal* stimuli were employed in the experiments described in (2) and (3) above, the increase in tension produced by repetitive stimulation cannot be ascribed to more nerve and muscle fibres

coming into action. The greater tension of a tetanus compared with a twitch is thus due to *each muscle fibre generating a greater tension when repetitively stimulated.*

These results are of the greatest practical importance. In both reflex acts and voluntary movements the behaviour of the motor units must depend on the character of the discharge from the motor neurones. It can be proved that their *rate of discharge* may vary from very low to very high levels, *i.e.* from 5–10 to 100–150 per second; the degree of tetanus resulting (*i.e.* whether partial or complete) and the consequent nature and strength of the contraction will vary correspondingly in the way detailed above. Further, there is some delay (slight, it is true) before the tetanus develops its maximum

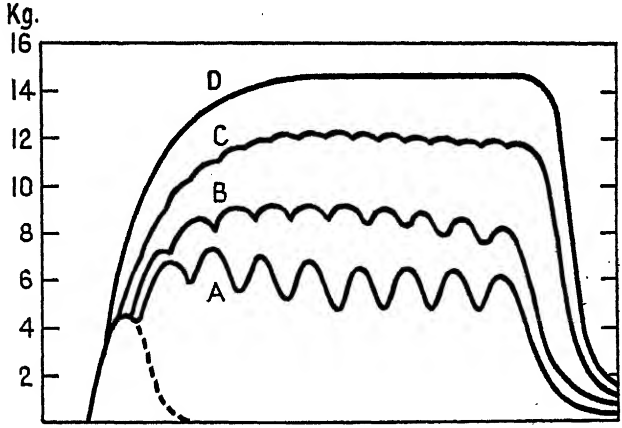


FIG. 300.—Genesis of Tetanus. (After Cooper and Eccles, *J. Physiol.*, 1930, *69.*)

Response of Mammalian Nerve-Muscle Preparation.
Isometric contraction records. Tension (ordinate) in kg.
Lowest curve—response of muscle to single *maximal* stimulus to motor nerve (simple twitch).
A, B, C, D—responses to rapidly interrupted *maximal* repetitive stimuli: A at 19, B at 24, C at 35, and D at 115 stimuli per second. Curves A, B, C show partial tetanus; curve D shows full tetanus.
Note as the frequency of stimulation rises the *tension developed becomes greater and is sustained more steadily.*

tension, so that the *duration* of the motor neurone discharge may be of importance also. The degree of activity of each motor unit can thus be *finely graded from the centre.* Further, the *number* of ventral horn cells activated during any reflex or other kind of act may be varied; the number of motor units in action at any moment is thus regulated, and obviously the larger the number of active motor units (all other factors being the same) the greater the tension resulting. Reference will be made to these principles again and again in subsequent sections (*e.g.* pp. 542, 585, 648).

There is another point of great importance in connection with the low rates of motor neurone discharge. It may be argued quite properly that the resulting partial tetanus, owing to its tremulous character, would be of little use either for maintaining positions or carrying out movements. This disadvantage is overcome in a subtle way by making the central discharge *asynchronous,*

i.e. the cells in what is called the *motor neurone pool*, which innervates the muscle, discharge out of step and do not fire off impulses simultaneously. As a result, the different muscle units are all (at any one moment) in different phases of activity; when one group is contracting another is relaxing, and vice versa. Algebraic summation occurs, the individual variations are evened out and the muscle gives a steady pull. This idea is represented diagrammatically for two motor units in Fig. 301. These views will be used to explain the steady pull in decerebrate rigidity (p. 585) and other postures, and in weak movements such as those of the diaphragm in quiet breathing (p. 392).

Muscle can contract under two sets of conditions: (i) *isotonic*, where free shortening is permitted and its degree can be recorded; (ii) *isometric*, where shortening is reduced to a minimum and the tension developed is measured.

Electrical Changes in Skeletal Muscle.—(1) Stimulation of a muscle through its nerve is called *indirect* stimulation.

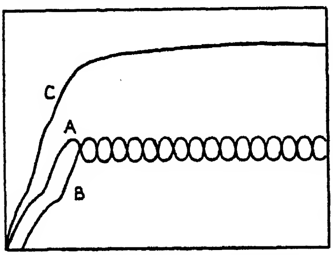


FIG. 301.—Asynchronous Discharge converting Partial Tetanus into a Steady Pull.

Two motor units (A and B) are represented, which are stimulated at the same low rate so that *each unit* responds with a partial tetanus and therefore with a tremulous contraction. If the interval between the onset of stimulation of A and B is suitably spaced (as explained in the text), the combined contraction of the two units leads not only to a greater tension but to a steadily maintained contraction (curve C).

(i) When a muscle is stimulated indirectly, a localized potential first appears at the motor end plate as explained on p. 512. When this end plate potential attains a critical magnitude it generates a propagated muscle action potential which travels simultaneously to both ends of the muscle fibre (Fig. 302).

(ii) A single stimulus produces a single muscle action potential; it is completed during the phase of rising mechanical tension of the twitch (Fig. 302, A).

(iii) If the nerve is stimulated *repetitively* a muscle action potential develops in response to *each* stimulus in the series even when the mechanical record shows complete fusion of the contraction waves (Fig. 302, B). The electrical record thus indicates the rate at which the nerve is being stimulated.

(iv) When an electrical stimulus is applied directly to a normal muscle, the response is due to stimulation either of the nerve fibres within the muscle or of the motor end plates; the stimulation employed is thus still indirect.

(2) *Direct* stimulation of the muscle fibres can be achieved after treatment of the muscle with curare which blocks neuromuscular transmission (p. 516).

(i) When a curarized muscle is stimulated directly the sequence of events resembles that occurring in a stimulated nerve fibre. A local catelectrotonus is first produced; when this reaches a critical magnitude a propagated action potential is generated (cf. nerve spike potential) which travels at a low velocity. *The interior of a resting muscle fibre is negative compared with the surface; during activity the surface becomes negative compared with the interior, i.e. the resting polarization is reversed* (as in nerve fibre (p. 484) and cardiac muscle (p. 240)). The muscle action potential is due to the movement of ions between the intracellular and extracellular fluids (cf. p. 485).

(ii) The shape of the action potential depends (as in nerve) on the recording

conditions. When one electrode is on an injured patch of muscle and the other is on intact muscle the wave is *monophasic*; when both surface electrodes are on normal muscle the record is usually *diphasic* (cf. p. 485).

(3) When a muscle is in a state of tone or is contracting voluntarily, the motor neurones (as explained on p. 585) discharge ("fire") *asynchronously*. The electrical record obtained under these conditions from the muscle *as a whole* (with belly-tendon or surface leads) shows rapid irregular variations, quite unlike the rhythmic, repetitive pattern of the electrical record in Fig. 302, B, which is the result of *synchronous* stimulation of all the motor nerve fibres. The belly-tendon or surface electrode leads record the *sum* of all the motor unit activity present at any moment in the muscle. In order to observe the rate of firing of *individual motor units* and, therefore, of *individual ventral*

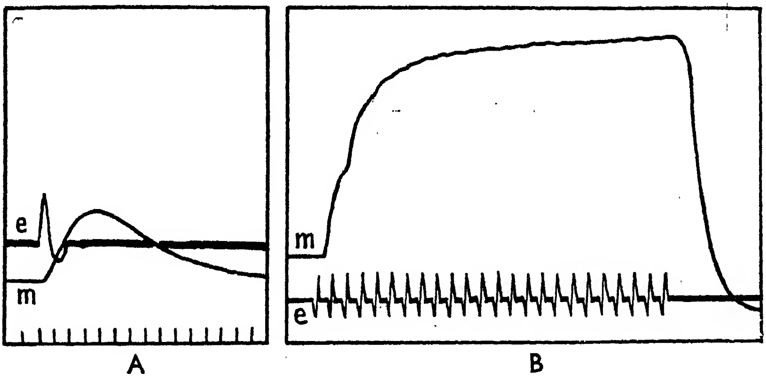


FIG. 302.—Mammalian Nerve-Muscle Preparation. Electrical and Mechanical Changes in Skeletal Muscle in response to Motor Nerve Stimulation. (Sherrington *et al.*, *Reflex Activity of Spinal Cord*, 1932.)

m = mechanical record. e = electrical record of whole muscle. One recording electrode is placed on the belly of the muscle, the other on the tendon (belly-tendon lead).
A. Response to single stimulus. Diphasic action potential which is completed in the early part of the contraction phase. Time in 0·01 second.
B. Response to stimulation at the rate of 67 per second. There is almost complete fusion of the mechanical contraction waves, but the action potential waves are distinct and discrete and follow the stimulation ate.

horn cells, it is necessary to use a concentric needle electrode, *i.e.* a hypodermic needle down which a fine insulated wire is inserted so that the bare tip of the wire just shows at the point of the needle. This limits the electrical pick-up to the muscle fibres in the vicinity of the needle tip, and it is then relatively easy to record the activity of individual motor units and to determine their discharge rates (Fig. 419).

(4) Although the motor unit is the physiological unit of muscle action, it is possible in cases of injury or disease of the lower motor neurone to record the action potentials resulting from the spontaneous random discharge of *single* muscle fibres or of groups of fibres constituting a *fraction* of a motor unit (*fibrillation potentials*, Fig. 303, A, B). The normal motor unit action potential is the sum of the action potentials of all the individual muscle fibres of which the unit is composed. Thus the motor unit action potential

is of greater amplitude than the single fibre action potential; it is also of longer duration because the muscle fibres of a motor unit do not fire off absolutely synchronously but with a temporal dispersion of some 5–10 milliseconds (Fig. 303, N).

FATIGUE.—If motor nerve stimulation is repeated for a sufficient length of time, fatigue develops. The latent period of the contraction becomes

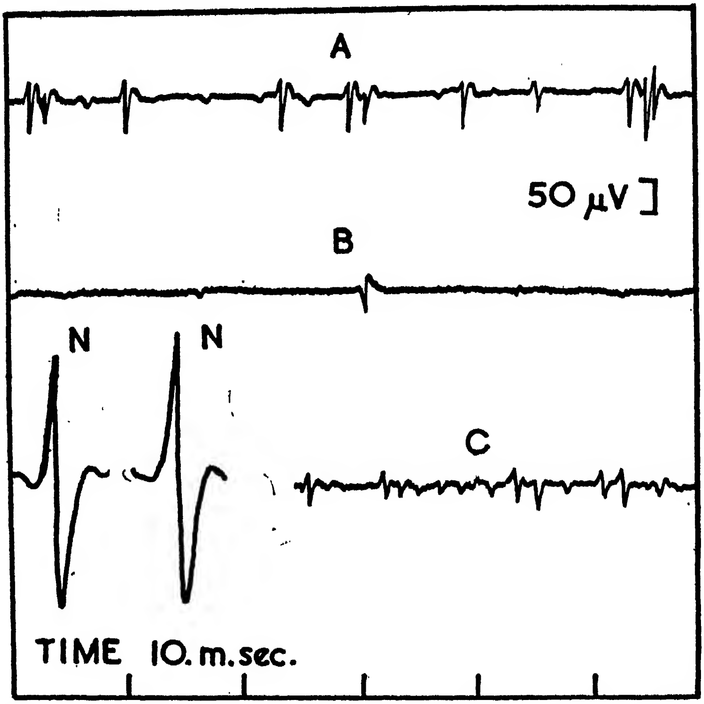


FIG. 303.—Normal Muscle Action Potentials and Fibrillation Potentials. (Weddell *et al.*, *Brain*, 1944, *67*.)

Records taken with concentric needle electrode in skeletal muscle in man.
N, N. Representative action potentials of motor unit of normal muscle. Potentials are mainly diphasic. Note their magnitude and duration.
A. Fibrillation potentials in denervated muscle. The deflections are small and very brief.
B. Occasional fibrillation potentials in another denervated muscle.
C. Same muscle as in B after administering prostigmine (an anticholinesterase). The frequency of the fibrillation potentials is greatly increased. Time in 10 m.sec.

longer, the rise of tension is slower and smaller, and relaxation is more gradual and incomplete. Finally, the muscular response ceases altogether. If the muscle is now stimulated *directly* it again responds well until finally fatigue once more sets in. The site of fatigue with indirect stimulation is obviously not in the muscle fibres themselves nor in the nerve fibres (which are for practical purposes unfatiguable); it is believed to depend on changes in the neighbourhood of the motor end plate, but their nature is unknown. The cause of fatigue of muscle to direct stimulation is also unknown.

All-or-None Relationship between Strength of Stimulus and Size of Response.—(i) When curarized skeletal muscle is stimulated directly, it displays the all-or-none relationship in the sense already defined for nerve (p. 489). The *individual* muscle fibre does not respond at all if the stimulus is too weak; it responds maximally (for the prevailing environmental conditions) when the stimulus rises to threshold; the contraction is not increased if the stimulus strength is further raised. Stronger stimuli, however, progressively bring *more muscle fibres* into action and thus the tension of a muscle increases as the strength of the stimulus applied to it rises.

(ii) Similarly when a motor nerve is stimulated the muscle fibres which it supplies always react to the best of their "ability." If a stronger stimulus is applied to the motor nerve *more nerve fibres* are activated and consequently *more groups of muscle fibres* contract, all of them to the best of their ability.

(iii) These results are of purely academic interest, because under natural conditions in the body, artificial stimuli of varying strength are not applied directly to the muscle or to its motor nerve fibres; the nerve fibres are activated by a *discharge from their parent cells.* More motor nerve fibres are brought into action in the body when *more ventral horn cells* discharge. It is very important also to appreciate that the responsiveness of the muscle is modified by many *local* conditions, and that both experimentally and naturally the same train of nerve impulses may lead to wide variations in resulting tension. Thus the mechanical response of a muscle to a maximal stimulus supplied to a motor nerve is greater with greater *initial length* of muscle fibre (within certain limits), or *higher temperature*; it is also affected by the freshness of the preparation and the adequacy of the blood and oxygen supplies (it is diminished by fatigue or asphyxia). The response to stimulation at high rates, the tetanus, as already emphasized, has *four or five times the tension of* the twitch. It will be explained later that the muscle response is greatly modified by *drugs* such as *curare* (p. 516) or *eserine* (cf. p. 515).

Effects of Section of Motor Nerve. Lower Motor Neurone Lesion.[1]—Injury or destruction of the ventral horn cells (or the cells in the motor cranial nuclei) or of the motor fibres supplying the muscles produces a characteristic series of changes.

(1) Results of Section of a Motor Nerve.—(i) The nerve fibres distal to the point of section undergo degeneration (p. 496); this applies to both the efferent fibres (and ultimately to the motor end plates, *infra*) and to the afferent fibres (and the muscle sense-organs from which they arise).

(ii) The ventral horn cells (and to a minor extent the cells of the dorsal root ganglia) undergo chromatolysis.

(iii) The muscle fibres which have been deprived of their efferent nerve supply become completely paralysed; all reflexes including reflex tone are abolished and so the muscles are flaccid.

(iv) After three months the motor end plates become distorted or disappear. The denervated muscle fibres progressively shrink, presumably from disuse, for a period up to three years. Later, if re-innervation has not taken place, "disruptive" changes occur. The muscle fibres split longitudinally into individual fibrils and also fragment transversely; later they are converted into tubes filled with deeply staining nuclei and granular material.

[1] Weddell *et al., Brain,* 1944, *67,* 178. Bowden and Gutmann, *ibid.,* 273. Ritchie, *ibid.,* 314.

is much longer than that of nerve fibres, the strength-duration curve is correspondingly altered. Fig. 304 shows that stronger stimuli or currents of longer duration must be employed to elicit a muscular contraction. Should the motor nerve fibres regenerate later, the curve gradually returns to the normal pattern (Fig. 305). When the denervated muscle fibres degenerate completely no response is obtained to electrical stimulation of any duration or (tolerated) intensity.[1]

(2) If the *ventral horn cells* are destroyed, no regeneration, and therefore, no recovery can occur. The functional loss is purely *motor* (and not sensory, too).

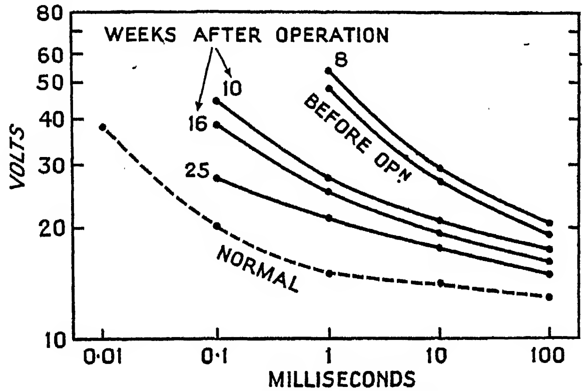


FIG. 305.—Effect on Strength-Duration Curve of Human Muscle of Section of Motor Nerve followed by Resuturing and Regeneration of Motor Nerve Fibres. (Ritchie, *Brain*, 1944, *67*, 322.)

Ordinate : Strength of stimulus in volts.
Abscissa : Duration of current flow in milliseconds

ELECTRICAL AND CHEMICAL TRANSMISSION OF THE NERVE IMPULSE

When a nervous impulse, travelling in an efferent peripheral nerve reaches the terminals of its fibre, it produces a characteristic response in the effector tissue, *e.g.* contraction of skeletal muscle, increased or decreased activity of smooth or cardiac muscle, secretion of glands or discharge of autonomic ganglion cells. The question arises: what is happening at the "junctional tissue" between the nerve terminals and the effector tissue, *e.g.* at the motor end plate, the synapses in autonomic ganglia or the autonomic postganglionic endings in smooth muscle and glands.

[1] *Reaction of Degeneration.*—The motor point is stimulated with (i) a galvanic current of fairly long duration, (ii) an interrupted faradic current of much shorter duration. Normally innervated muscle responds to either form of stimulation. When the motor nerve fibres have completely degenerated, the denervated (but intact) muscle fibres owing to their longer chronaxie do not respond to the faradic stimulus (as the duration of the current flow is too short); but the longer current flow of the galvanic stimulus is still effective if it is of sufficient strength. The loss of response to faradic stimulation with persistence of response to galvanic stimulation is called the reaction of degeneration. When the muscle fibres degenerate too, all electrical responses are abolished.

Two main views have been put forward :

(1) The transmission process at nerve ends (like that responsible for conduction in the nerve fibre) is essentially *electrical*. When the spike potential reaches the nerve ends it *directly* sets up a localized *catelectrotonus* ; when the latter reaches a critical magnitude it depolarizes the surface membrane of the adjacent tissue and thus stimulates it.

Such a catelectrotonus has been demonstrated at the motor end plate (*end plate potential*) and at the synapses in autonomic ganglia (*synaptic potential*). A state of catelectrotonus may be associated with excitatory transmission at synapses in the central nervous system. The nature of the electrical change at autonomic postganglionic terminals is obscure.

(2) The alternative view is that a chemical intermediary or *chemical transmitter* intervenes between the spike potential (nerve impulse) and the processes which stimulate the effector tissue. Chemical transmitters have been demonstrated : (i) at the terminals of all postganglionic fibres (sympathetic and parasympathetic) ; (ii) probably at the terminals of the vasodilator fibres in the dorsal nerve roots ; (iii) intervening between the spike potential and the end plate potential at motor end plates ; (iv) probably intervening between the spike potential and the synaptic potential in autonomic ganglia. (v) It has been argued that if chemical transmission occurs at the synapses in autonomic ganglia similar processes may be involved in transmission at the synapses in the central nervous system (p. 528).

Nature of Transmitters.—Two chemical transmitters have been identified : (i) acetylcholine ; (ii) a substance which resembles adrenaline closely both chemically and in its pharmacological properties ; the cautious call this transmitter " sympathin " or " adrenaline-like." There is evidence that it may be a mixture of adrenaline and nor-adrenaline (adrenaline minus its terminal methyl group) in varying proportions (p. 722). In this account, this transmitter will for simplicity be called adrenaline.

Cholinergic and Adrenergic Fibres.—Fibres which release acetylcholine at their terminals are called cholinergic (*i.e.* acting through the mediation of acetylcholine) ; fibres which release adrenaline at their terminals are called adrenergic.[1] In the peripheral nervous system, the only known adrenergic fibres are the postganglionic fibres of the sympathetic nervous system. The cholinergic fibres are of much wider distribution : they include the postganglionic fibres of the parasympathetic system ; all the preganglionic fibres of the sympathetic and parasympathetic systems, *i.e.* those ending in autonomic ganglia ; the motor fibres to skeletal muscle ; probably the antidromic vasodilator fibres in the dorsal nerve roots supplying skeletal muscle. In addition there are some postganglionic fibres which though they belong to the anatomical sympathetic system, are cholinergic, *e.g.* the sympathetic supply to the sweat glands.

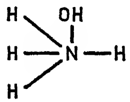
Formation, Release and Destruction of Transmitters.[2]—Acetylcholine.—All cholinergic nerves contain the enzymes necessary for the synthesis of acetylcholine[3] ; acetylcholine can be extracted from these nerves.

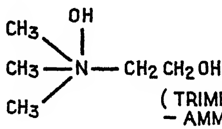
[1] *Histaminergic* fibres have also been recognized which release histamine at their terminals. They are found in the cutaneous branches of the dorsal nerve roots (p. 309).

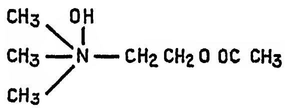
[2] Burn, *Physiol. Rev.*, 1950, *30*, 177. (Discussion of local action of acetylcholine, adrenaline and histamine.)

[3] The enzyme is called *choline acetylase.*

It is thought that acetylcholine is synthesized within each cholinergic neurone, probably mainly in the cell body; it is certainly diffused throughout the length of the axon. The arrival of the impulse at the peripheral nerve terminal alters the local permeability and so enables a small amount of acetylcholine to leak out, *i.e.* to be released from the fibre, and come into contact with the effector tissue. At cholinergic nerve terminals a specific enzyme, *cholinesterase*,[1] is found in high concentration; it hydrolyzes acetylcholine to choline, which is comparatively inactive, and acetic acid. In this way the concentration of acetylcholine is rapidly reduced below threshold level thus preventing continued excitation. The chemical constitution of acetylcholine and related substances is shown below (cf. p. 867):

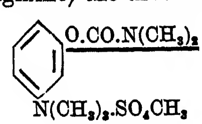

AMMONIUM HYDROXIDE


CHOLINE
(TRIMETHYL-HYDROXY ETHYL - AMMONIUM HYDROXIDE)

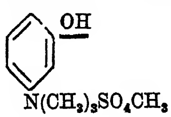

ACETYLCHOLINE

Anticholinesterases.[2]—Three quite distinct classes of chemical substance have a powerful inhibitory effect on cholinesterase.

(i) Compounds containing a urethane side chain, like eserine and prostigmine. The urethane can be represented as –O–OC–N(R_1, R_2) (where R_1, or R_2 or both, are alkyl or phenyl groups, attached to a nitrogen grouping); the rest of the molecular configuration may vary considerably. In Formula A, (prostigmine) the urethane grouping has been underlined. If the urethane


Prostigmine

A


m-hydroxyphenyl trimethyl ammonium methyl sulphate

B

grouping is replaced by an OH grouping (underlined) as in Formula B, the anticholinesterase activity disappears.

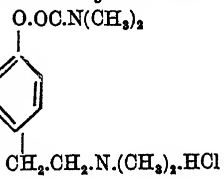
These anticholinesterases act by the method of *substrate competition*. The hydrolysis of acetylcholine by cholinesterase depends on a preliminary union of the two substances; the anticholinesterases also unite with cholinesterase and to that extent prevent the latter from uniting with, and acting on, acetylcholine.

[1] Whittaker, *Physiol. Rev.*, 1951, *31*, 312.
[2] Koelle and Gilman, *Pharmacol. Rev.*, 1949, *1*, 166

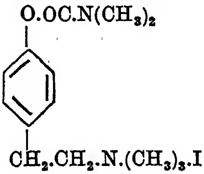
This group of anticholinesterases can be further subdivided into *tertiary* and *quaternary* compounds. Expressed most simply, the tertiary compounds are hydrochlorides or sulphates; the quaternary compounds are methiodides or methyl sulphates. The former are salts of tertiary ammonium bases; the latter have their basic nitrogen atom in the form of quaternary ammonium groups. *Eserine sulphate* is a typical tertiary compound; *prostigmine* (which is a methyl sulphate) is a typical quaternary compound.

The graphic formulæ set out below for two anticholinesterases, dimethyl carbamic hordenine HCl (tertiary) and dimethyl carbamic hordenine methyl iodide (quaternary), show these differences well.

Typical Tertiary Compound:
Dimethyl carbamic ester of hordenine *hydrochloride*



Typical Quaternary Compound:
Dimethyl carbamic ester of hordenine *methyl iodide*



(ii) Di-isopropylfluorophosphonates (DFP) and related substances.

(iii) Simple polyphosphates: tetra-ethyl pyrophosphate and hexa-ethyl tetraphosphate (Fig. 311A, B).

Adrenaline.—Extracts of adrenergic fibres contain adrenaline (and nor-adrenaline). The adrenergic neurones (like the adrenal gland cells) presumably synthesize adrenaline which is released at the postganglionic nerve terminals. It is thought that an enzyme called amine oxidase, rapidly destroys adrenaline and so prevents its undue persistence at the endings. It is claimed that the drug *ephedrine* inhibits the action of amine oxidase and thus prevents the adrenaline from being destroyed.

The pharmacological actions of acetylcholine are described on p. 718, and those of adrenaline on p. 724.

NEUROMUSCULAR TRANSMISSION[1]

The motor fibre nerve, on reaching the muscle fibre, loses its medullary sheath; the naked axis cylinder ramifies in a patch of granular nucleated muscle cytoplasm called the *muscle sole* (Fig. 306). The neuromuscular junctional region, is the *motor end plate.* The nerve filaments and the muscle tissue make extensive *contact* with one another but there is no *continuity* of tissue; the specialized axon membrane separates the two tissues. When the nerve impulse (signalled by the spike) reaches the end plate it releases a

[1] Dale, Feldberg and Vogt, *J. Physiol.*, 1936, *86*, 353; Brown, *ibid.*, 1937, *89*, 220 Eccles *et al.*, *J. Neurophysiol.*, 1941, *4*, 362, 402; 1949, *12*, 59. Kuffler, *ibid.*, 1942, *5*, 18' 309; 1943, *6*, 99. Nachmansohn, in *The Hormones* (ed. Pincus and Thimann) N.Y., 1950, *2*' Hunt and Kuffler, *Pharmacol. Rev.*, 1950, *2*, 96. Feldberg, *Brit. med. J.*, 1951, i, 967.

minute dose of a *chemical transmitter*, namely acetylcholine, which either directly, or, more probably, indirectly through some further intermediary. "depolarizes" the end plate producing a *localized non-propagated, electrotonic potential* known as the *end plate potential*. When the end plate potential

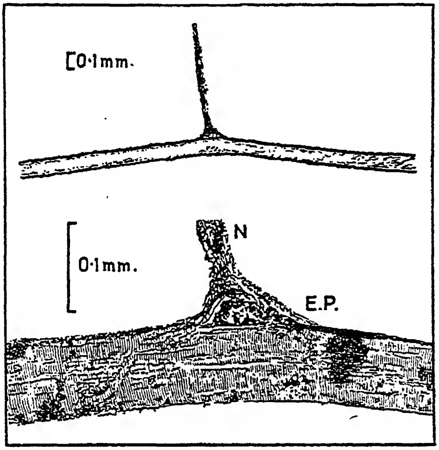


FIG. 306.—Motor End Plate. (Kuffler, *J. Neurophysiol.*, 1942, *5*, 197.)

Photograph of living single nerve fibres (N) ending in a motor end plate (E.P.) on a single skeletal muscle fibre (frog). Upper figure is at smaller magnification than lower figure.

attains a certain critical magnitude, it "depolarizes" the surface membrane of the muscle fibre setting up a *propagated* muscle action potential. The sequence of events is shown in the Table below.

Muscle Fibre Stimulated through its Motor Nerve

Nerve Fibre impulse (spike potential).

↓

Impulse reaches end plate

↓

Release of acetylcholine

↓

(?) Other consequential changes

↓

Local end plate potential

↓

Propagated muscle spike potential→Complicated, ill-understood changes

↑

Development of muscle tension (contraction)

The appearance of the muscle spike potential just precedes the development of mechanical tension (Fig. 300). The causal relationship, if any, between these two events is not understood. Current views on the way in which muscular tension develops are discussed on p. 427 but they take no account of the muscle spike potential except in so far as it normally heralds the onset of physico-chemical changes in the muscle fibre. Certain types of

contraction are known which are *not* accompanied by an action potential; they are referred to as *contractures*.

End Plate Potential.—The most instructive preparation for the study of the end plate potential is one consisting of a *single* motor nerve fibre and a *single* muscle fibre (Fig. 306); the proximal recording electrode is placed on the end plate and the distal electrode on a point further away on the muscle fibre.

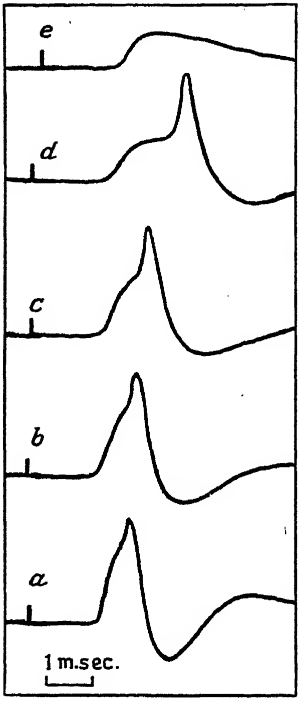


FIG. 307.—Effect of Progressive Curarization on End Plate Potential. (Kuffler, *J. Neurophysiol.*, 1942, *5*, 23.)

Experiment on single nerve-muscle fibre preparation (as in Fig. 306). Proximal electrode on end plate; distal electrode on muscle fibre.
At signal (short vertical line) stimulate motor nerve fibre.
e: End plate deeply curarized. End plate potential only. No propagated potential.
d, c, b: Progressively milder curarization.
a: No curare (cf. Fig. 308, *a, e*).

(1) The preparation is treated with *curare* in a concentration which abolishes the propagated muscle potential, but not the end plate potential. Under these conditions (Fig. 307, *e*) stimulation of the motor nerve fibre produces a slowly rising and more slowly declining negativity at the end plate; the deflection is *monophasic*, indicating that the disturbance is *localized* and not propagated; it is the end plate potential (in a reduced form) which has *not* been followed by a propagated muscle fibre potential. The potential is maximal at the end plate and rapidly diminishes in magnitude (decrements) with passage of time and with distance from the end plate (it disappears a few mm. away).

With progressively *weaker* concentrations of curare, the response is modified as follows:

(i) A sharp *spike* arises from the plateau of the end plate potential; this represents the initial deflection of a propagated muscle potential as proved by the fact that it is followed by a final downward deflection. The initial spike and the final deflection together constitute the diphasic muscle fibre action potential (Fig. 307, *d*).

(ii) With still less curare the end plate potential rises more *rapidly* to a *greater height* and is followed after a progressively *shorter interval* by the typical propagated muscle potential (Fig. 307, *c, b*).

(2) The response of an *untreated single* end plate can now be studied. In Fig. 308 the proximal electrode was applied at varying distances from the end plate; the distance in *a* was very small, 80 μ in *b* and 230 μ in *c*. The initial upward deflection in each case is the end plate potential; because of typical decrement with distance it is smaller and more slowly rising in *c* and *b* than in *a*. The end plate potential is followed in each instance by a propagated (diphasic) spike potential. In Fig. 308, *e*, the proximal electrode was applied

exactly on the end plate ; this enables the true course of the normal end plate potential to be followed. It is seen that the end plate potential rises rapidly to the full height of the spike of the propagated muscle action potential which it generates and then merges with it ; the fact that a propagated spike has been produced is proved by the succeeding descending deflection.

(3) SUMMATION OF END PLATE POTENTIAL.—Under deep curarization (as already mentioned) the end plate potential is reduced and is not followed by a propagated potential. If a second impulse arrives at the end plate before the previous end plate potential has died away, a second potential is superimposed on the first (Fig. 309). The second potential is always bigger than the first when measured from its point of origin, *i.e.* the presence of the first end plate potential facilitates the development of a subsequent potential. If the two or more nerve stimuli are suitably spaced, the end plate potential is progressively built up, *i.e.* it undergoes summation, till it becomes big enough to produce a propagated disturbance, and the muscle responds. The properties of the curarized end plate, especially the phenomena of facilitation and summation which it displays, are of great interest because they resemble in several respects the properties of synapses in ganglia and in the central nervous system (cf. p. 524 ; p. 533).

The rôle of acetylcholine in the production of the end plate potential is described below.

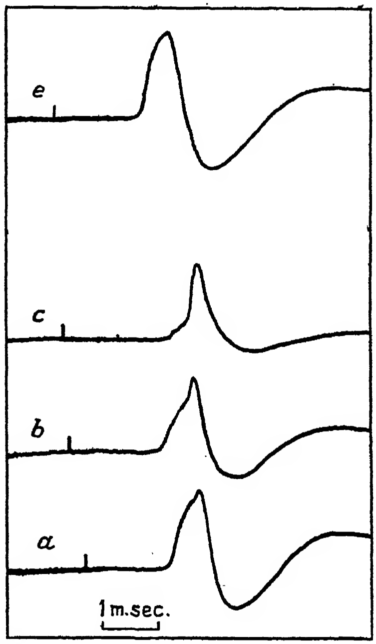


FIG. 308.—Decrement of End Plate Potential with distance from End Plate. (Kuffler, *J. Neurophysiol*, 1942, *5*, 21.)

The preparation was *not* curarized.
Position of *proximal* electrode is as follows :
a : slightly away from end plate (of Fig. 307, *a*).
b : 80 μ distance from end plate.
c : 230 μ distance from end plate.
e : exactly at end plate.
The distal electrode was placed farther away on the muscle fibre.

Acetylcholine as Chemical Transmitter at the Motor End Plate.—The principal experimental evidence is as follows :

(1) RELEASE OF ACETYLCHOLINE IN INDIRECTLY STIMULATED MUSCLE.—The cat's tongue or hind limb is perfused with fluid containing eserine (to inhibit the action of cholinesterase and thus prevent the destruction of any acetylcholine that may be released (cf. p. 509)). Stimulation of the hypoglossal nerve or the motor nerves to the leg, leads to the appearance of acetylcholine which can be demonstrated in the outflow fluid by appropriate tests (p. 719).

(2) ACETYLCHOLINE PRODUCES PROPAGATED MUSCLE ACTION POTENTIALS AND MUSCULAR CONTRACTION.—It is necessary to demonstrate that the released acetylcholine is causally related to muscular excitation and contraction.

(i) If acetylcholine in minute doses is injected into the blood vessels supplying a muscle (so-called close arterial injection) to obtain rapid access to the muscle fibres, it elicits a brief contraction of normal mammalian muscle *in situ* and with intact circulation. As small a dose of acetylcholine as 5 μg. (=0·005 mg.) may set up a contraction of higher tension than that resulting from maximal stimulation of the motor nerve with a single electrical shock. The muscular response to injected acetylcholine is not a simple twitch accompanied by a single spike potential; it is a brief *tetanus*, the contraction being accompanied by asynchronous repetitive action potentials arising in many motor units which have been stimulated to short bursts of activity, but which are out of phase with one another.

(ii) When a minute dose of acetylcholine is applied directly to the end

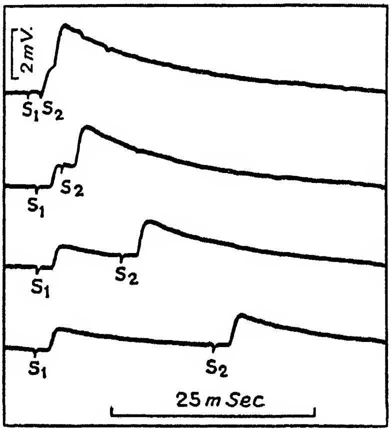


FIG. 309.—Summation of End Plate Potentials in Curarized Nerve-Muscle Preparation. (Eccles, Katz, and Kuffler, *J. Neurophysiol.*, 1941, *4*, 377.)

Small downstroke: application of stimulus (S_1 or S_2) to the motor nerve.
Note the summation of end plate potential.

plate region, it sets up a small end plate potential which rises and declines slowly but remains localized; higher concentrations set up a bigger and more rapidly rising end plate potential which on attaining threshold value sets up a propagated muscle potential (Fig. 310). Still higher concentrations produce *repetitive* responses from the muscle, the frequency of which goes up with increasing concentrations of acetylcholine. Acetylcholine, applied elsewhere along the course of the muscle fibre, produces no potential changes.

The muscle action potentials set up by injected or applied acetylcholine are identical in character with those occurring in the muscle on stimulation of the motor nerve.

(3) RÔLE OF CHOLINESTERASE.—This enzyme, which rapidly destroys acetylcholine, is found in high concentration in the motor end plate region (but in much smaller amounts elsewhere along the course of the muscle

fibre). Its presence at this site ensures that the acetylcholine normally released in response to each nerve impulse only gives rise to a *single* response of the muscle fibres and does not persist long enough or in sufficient concentration to produce the repetitive (subtetanic) responses seen when acetylcholine is injected intra-arterially.

(4) EFFECTS OF ANTICHOLINESTERASES.—Anticholinesterases inhibit the action of cholinesterase and thus preserve acetylcholine which has been released naturally or has been injected into the circulation (cf. p. 509). Following the injection of one of these drugs, the response of skeletal muscle to maximal single shocks applied to the motor nerve is considerably enhanced (or *potentiated*) and prolonged (Fig. 311A). Electrical investigation shows that under these conditions each nerve volley reaching the muscle fibres produces not a simple twitch, but a *repetitive* response, *i.e.* a short tetanus which of course produces a greater tension than the twitch; thus Fig. 311B shows *sustained* electrical activity following a single nerve stimulus. The anticholinesterase preserves the acetylcholine naturally released at the motor end plate for long enough above threshold concentration to enable it to produce repetitive responses analogous to those illustrated by Fig. 310.

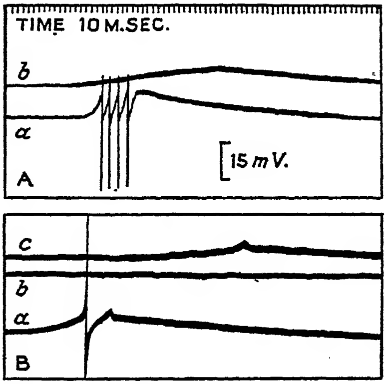


FIG. 310.—Effects of applying Acetylcholine directly to Motor End Plate of Single Nerve Fibre-Muscle Fibre Preparation. (Kuffler, *J. Neurophysiol.*, 1943, *6*, 102.)

Proximal electrode is at point of application of solution.

A. (*a*) Acetylcholine in concentration of 10^{-6} applied to end plate sets up four propagated (diphasic) muscle potentials.

(*b*) Weaker acetylcholine concentration produces small localized negative (monophasic) end plate potentials only.

B. (*c*) Acetylcholine in concentration of 10^{-6} produces localized (monophasic) end plate potential only.

(*b*) Application of saline: no effect.

(*a*) Acetylcholine in stronger concentration than 10^{-6} produces a single (diphasic) propagated muscle fibre potential followed by a localized (monophasic) end plate potential.

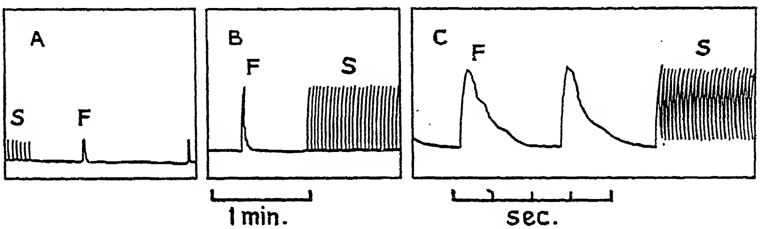


FIG. 311A.—Effect of Anticholinesterase on Mechanical and Electrical Response of Nerve-Muscle Preparation. (Chennells, Floyd and Wright, *J. Physiol.*. 1949, *108*, 375) (cf. Fig. 311B).

A. Cat. Chloralose. Isometric myogram of quadriceps muscle in response to maximal stimulation of motor nerve at rate of one in 2 sec.

A. Control.

B, C. After injection of anticholinesterase (tetraethyl pyrophosphate). The muscle response is enhanced (potentiated) and more prolonged (it is now a short *tetanus* not a twitch.)

S. Record on slow drum (time scale, in minutes).

F. Record on fast drum (time scale, in seconds).

(5) Response of Denervated Muscle.—When the motor nerve is cut and time allowed for it to degenerate completely, direct stimulation of the muscle elicits a response which is *not* accompanied by a release of acetylcholine proving that the transmitter is released only at the nerve endings in the motor end plate. Such denervated muscle is, however, much more sensitive to the effects of *injected* acetylcholine. This is an example of the more general phenomenon that denervated structures are more susceptible to the action of their natural transmitters. Thus the normal pupil does not dilate when adrenaline is introduced into the conjunctival sac; a good response is, however, obtained if the superior cervical ganglion is previously extirpated and the post-ganglionic fibres to the pupil are allowed to degenerate (cf. pp. 721, 729).

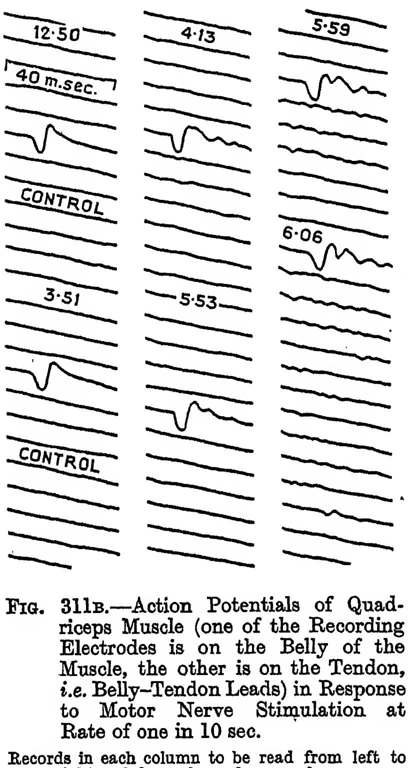


Fig. 311b.—Action Potentials of Quadriceps Muscle (one of the Recording Electrodes is on the Belly of the Muscle, the other is on the Tendon, *i.e.* Belly–Tendon Leads) in Response to Motor Nerve Stimulation at Rate of one in 10 sec.

Records in each column to be read from left to right and from above downwards.
Column 12·50, 3·51. Control. A simple diphasic variation develops in response to each stimulus.
Anticholinesterase (hexaethyl tetraphosphate) then injected. Note at 4·13 and 5·53 that a single stimulus applied to the nerve produces a diphasic variation which is followed by *repetitive after-activity*.
At 5·59 and 6·06 this after-effect is very prolonged (cf. Fig. 311a).

(6) Release of Acetylcholine.—It has been suggested that the nerve impulse may release acetylcholine from a precursor at the end plate; but there is no evidence to support this view. As explained on p. 508, cholinergic neurones can synthesize acetylcholine, which is diffused throughout the length of the fibre. When the impulse arrives at the nerve ending, it increases the permeability of the nerve fibre membrane locally, thus enabling some acetylcholine to leak out from the nerve terminals into the "muscle sole."

(7) Action of Curare.—It has been known since the experiments of Claude Bernard that curare progressively diminishes and finally abolishes the response of a muscle to indirect stimulation (*i.e.* to stimulation of the motor nerve) (Fig. 312). At this stage, however, *direct* stimulation of the muscle still produces a muscular contraction. As curare has no effect on the nerve fibre, Bernard rightly concluded that curare acted at the neuromuscular junction, *i.e.* in the region of the motor end plate. After administration of paralysing doses of curare, nerve stimulation still results in a normal release of acetylcholine; thus curare does not interfere with the release of the transmitter but prevents the released acetylcholine from producing a contraction. Similarly after curare, intra-arterially injected doses of acetylcholine which were previously effective, fail to cause a contraction (cf. p. 720).

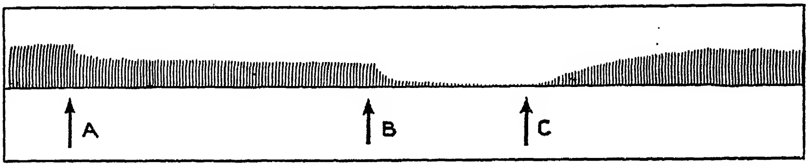


FIG. 312.—Action of Curare on Response of Nerve-Muscle Preparation. Anti-Curare Action of Anticholinesterase. (Chennells, Floyd and Wright, *J. Physiol.*, 1949, *108*, 384.)

Isometric myogram of quadriceps (cat). Femoral nerve maximally stimulated at rate of one in 2 sec
At A, B. Inject curare preparation; contraction height diminished.
At C. Inject anticholinesterase; recovery of contraction height.

The experiments illustrated by Fig. 307 demonstrate that curare interferes with the formation of the end plate potential by the released acetylcholine. As the end plate potential formed is too small to set up a propagated action potential no muscular contraction results.

(8) ANTI-CURARE DRUGS.—(i) The "blocking" action of curare at the end plate region can be overcome by administering an *anticholinesterase* (Fig. 312). The acetylcholine normally released is then preserved with resulting *repetitive* stimulation of the end plate. There is consequently progressive summation of the initially sub-threshold end plate potential (as in Fig. 309) till it attains a sufficient magnitude to set up a propagated muscle fibre potential and a contraction.

(ii) Potassium salts (Fig. 314) and adrenaline, ephedrine and related substances (Fig. 317) also exert an anti-curare action.

(9) ACTION OF BOTULINUS TOXIN.—The toxin of the botulinus bacillus paralyses the muscle by a peripheral action. In some unknown way, the toxin prevents the release of acetylcholine at the end plate.

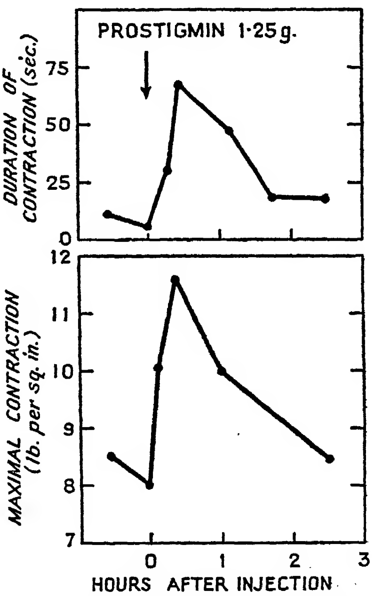


FIG. 313.—Effect of Anticholinesterase (Prostigmine) on Muscle Power in Case of Myasthenia Gravis. (Burgen, Keele, and McAlpine, *Lancet*, 1948, i, 519.)

At arrow inject 1·25 g. of prostigmine.
Upper record: maximal duration of contraction of muscles of hand and forearm.
Lower record: Maximal power of contraction of these muscles.

Myasthenia Gravis.[1]—Myasthenia gravis is a rare disease, characterized by great muscular weakness and rapid onset of fatigue, but without any recognizable changes in the nervous system or in the muscles themselves. The muscles first and most affected are the external ocular, the facial, and those concerned with mastication and swallowing.

[1] Viets and Schwab, *J. Amer. med. Assoc.*, 1939, *113*, 559. Walker, *Lancet*, 1934, i, 1200.

The muscular disability is probably due to some disturbance of the normal mechanism of transmission of the impulse at the end plate to the muscle fibres. Among the factors that have been examined are: (i) insufficient release of acetylcholine; (ii) abnormally rapid destruction of acetylcholine due perhaps to excessive local cholinesterase activity; (iii) presence of an abnormal curare-like substance hampering access to the muscle fibres of normally released acetylcholine. The evidence indicates that (iii) may be a causal factor.[1] In patients a substance is present in the blood leaving the muscles (especially active muscles) which depresses the response of an isolated nerve-muscle preparation to motor nerve stimulation. The following experiment shows that similar depression occurs in the patient: the arm muscles are exercised for about 4 minutes with their circulation completely occluded (by means of cuffs on the upper arms inflated to 200 mm. Hg). When the occlusion is released, *muscles elsewhere in the patient* show increased

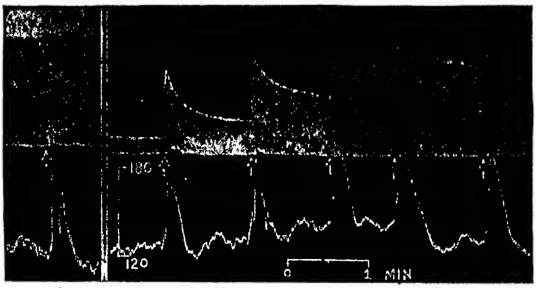


FIG. 314.—Anti-curare Action of Potassium Salts. (Wilson and Wright, *Quart. J. exp. Physiol.*, 1936, *26*.)

Isometric records of response of gastrocnemius (cat) to motor nerve stimulation (at intervals of 10 seconds) and blood pressure.

A. Normal muscle response. At arrow inject 30 mg. of KCl (as 3% solution) intra-arterially. A slight contracture develops.

B. Between A and B inject curare to reduce the response to nerve stimulation substantially. At each arrow inject again 30 mg. of KCl (3% solution). Note in each case the immediate increase in the muscle response, which subsequently declines.

weakness after a latent period of 10 seconds to 4 minutes; thus the upper lids show increased drooping (ptosis) and the facial muscles become feebler. It is noteworthy too that the distribution and character of the weakness in myasthenia gravis closely resemble those seen after injection of small doses of curare in man.

Administration of a suitable anticholinesterase, *e.g.* prostigmine or tetraethyl pyrophosphate (which protects acetylcholine from destruction) produces very striking clinical improvement: when so treated, patients with marked facial weakness and unable to swallow or get out of bed, rapidly develop a striking increase in muscular power (Fig. 313). They are able to eat a hearty meal, walk or run around the room, and perform light manual work. The beneficial effects last for a number of hours, depending on the drug and the size of the dose used. The anticholinesterase chosen should be one with

[1] Walker, *Proc. roy. Soc. Med.*, 1938, *31*, 722. Wilson and Stoner, *Quart. J. Med.*, 1944, *13*, 1.

a selective action on end plates; its peripheral action at parasympathetic ends (which would give rise to undesirable side actions such as slowing of the heart, fall of blood pressure, violent and painful intestinal contractions) is kept in check by means of atropine.

Certain other drugs with an anti-curare action have proved beneficial in myasthenia gravis though they do not antagonize cholinesterase. Among the more important are *potassium salts* (*infra*), *adrenaline* (p. 522), and *ephedrine* (p. 511). The beneficial action in myasthenia of so many diverse substances with only their anti-curare action in common supports the suggestion already made that in myasthenia some metabolic disorder may lead to the liberation of a curarizing agent.

Rôle of Thymus.—In many case of myasthenia gravis the *thymus* gland is enlarged. Surgical removal of the gland has been carried out in some patients, often with considerable benefit in respect of muscular strength: a few patients have recovered completely, but the majority required continued prostigmine therapy, though smaller doses of the drug were needed. It has been suggested (without adequate evidence) that the thymus may play an important part in the production of myasthenia gravis by liberating a curare-like substance.

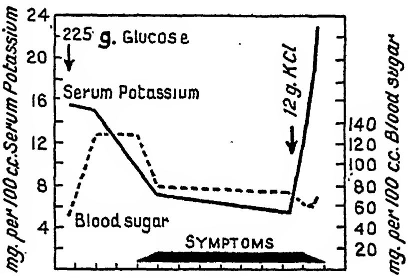


FIG. 315.—Blood Changes in Attack of Familial Periodic Paralysis. (Aitken, Allott, Castleden and Walker, *Clin. Sci.*, 1937, *3*.)

Blood sugar (dotted line) and serum potassium (continuous line) during an attack of paralysis induced by taking 225 g. of glucose (first arrow). The black band along the abscissa indicates the appearance and development of symptoms; they coincide with the decline in serum K^+, and are relieved by the administration of 12 g. of KCl which raises the serum K^+. Time (on abscissa) in 10 minute intervals.

Rôle of Potassium in Neuromuscular Transmission and Muscular Contraction.[1]—(1) Intra-arterial injection of potassium salts (*e.g.* 100 mg. in the form of a 1% solution of KCl) produces a powerful muscular contraction accompanied by normally propagated, repetitive action potentials, *i.e.* a short tetanus results. The response is indistinguishable electrically from that produced by acetylcholine. Potassium ions act selectively on the motor end plate. It has been suggested that acetylcholine, normally released or injected, somehow liberates potassium ions which are directly responsible for the production of the end plate potential.

(2) Solutions of ionized potassium salts have an anti-curarizing action. Fig. 314 shows that muscle responses to indirect stimulation which have been greatly reduced by the action of curare are rapidly restored to almost normal levels by repeated injection of KCl solutions.

(3) FAMILIAL PERIODIC PARALYSIS.[2]—This rare disease is characterized by attacks of paralysis which come on at irregular intervals. The fault lies peripherally: during an attack, stimulation of the motor nerve trunks or of the skin over the muscle elicits no response. The muscle thus reacts neither

[1] Wilson and Wright, *Quart. J. exp. Physiol.*, 1936, *26*, 127. Brown, *J. Physiol.*, 1937, *91*, 4P. Fenn, *Physiol. Rev.*, 1940, *20*, 377.

[2] Gass *et al.*, *Medicine*, 1948, *27*, 105.

to indirect nor direct stimulation. The attacks can be artificially produced in subjects of the disease by administration of large amounts of glucose, which are known to lead to a *marked fall of serum potassium* (*e.g.* from the normal level of 16–20 mg–% to about 10–12 mg–%). Symptoms of paralysis develop when the serum K^+ falls below 12 mg–% (Fig. 315). The rise of blood sugar itself is not the causal agent, because attacks can also be produced by an injection of insulin which likewise lowers the serum K^+, but of course also *lowers* the blood sugar. The administration of KCl (*e.g.* 12 g.) causes recovery, beginning in 10 minutes and becoming complete in about 7 hours, and is associated with restoration of the serum K^+ level.[1]

The evidence as a whole suggests that K^+ may have a dual rôle: (i) it may be concerned with neuromuscular transmission; (ii) the excitability and other properties of muscle may depend to some extent on the relative concentrations of K^+ within and without the fibres.

Myotonia Congenita.[2]—In this rare congenital disease, when the patient starts to perform a voluntary movement, the muscles are thrown into a

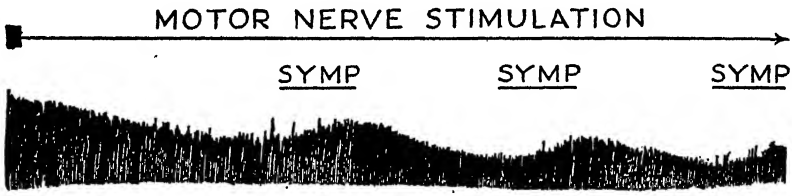


FIG. 316.—Effect of Sympathetic Stimulation on Fatigued Skeletal Muscle of Frog.

A series of rapidly recurring twitches were produced by stimulating the motor roots till signs of fatigue appeared. The sympathetic nerves were stimulated for the periods indicated by the signal above. After considerable latency the twitches increase in size, the maximum effect occurring after cessation of sympathetic stimulation. (After Orbell, *Pavlov Jubilee Volume*, 1924.)

state of spasm which gradually passes off; thus if he yawns the mouth remains open for some time. As the movement is repeated the muscles become more supple. If the muscle is stimulated electrically, contraction persists for a considerable time *after stimulation is discontinued*. An interesting strain of goat has been discovered in the U.S.A. suffering from a condition indistinguishable from human myotonia; as a result the condition can be examined experimentally. The muscles are extremely sensitive to mechanical stimulation; tapping, for instance, produces a long lasting contraction which is due to a long lasting tetanus in groups of muscle fibres, accompanied by normal action potentials. The response is obtained after cutting the motor nerve. The response of the muscle to a single motor nerve stimulus is not a twitch but is repetitive, and the tension is correspondingly greater than in normal animals. At the end of a bout of stimulation at a rate of 50 per second (to produce a tetanus), the muscle remains tetanically contracted as it does after sudden exertion in this disease. The abnormal sensitivity to mechanical stimulation is unaffected by curare or complete degeneration of the motor nerve ending; it is, however, progressively reduced by bouts of

[1] Aitken *et al.*, *Clin. Sci.*, 1937, *3*, 47. Allott and McArdle, *ibid.*, 1938, *3*, 229.

[2] Brown and Harvey, *J. Physiol.*, 1939, *96*, 11P. Kolb, *Johns Hopk. Hosp. Bull.*, 1939, *63*, 221. Denny-Brown and Nevin, *Brain*, 1941, *64*, 1.

stimulation at low rates, *e.g.* 5 per second. Sensitivity to intra-arterially injected acetylcholine is unchanged, but the duration of the response is increased. Sensitivity to injected potassium salts is however far greater. Myotonia is probably not due to an abnormality of neuromuscular transmission but to changes in the excitability of the muscle fibres themselves.

Quinine has proved helpful clinically; and it similarly relieves the myotonic condition in the goat.

Action of Sympathetic on Skeletal Muscle.—(i) Non-medullated sympathetic postganglionic fibres pass to the blood vessels of skeletal muscle (p. 308). It has been claimed that these nerve fibres also end on the surface of the sarcolemma of the muscle fibres; these claims are probably incorrect.

(ii) The sympathetic innervation is not responsible for the maintenance

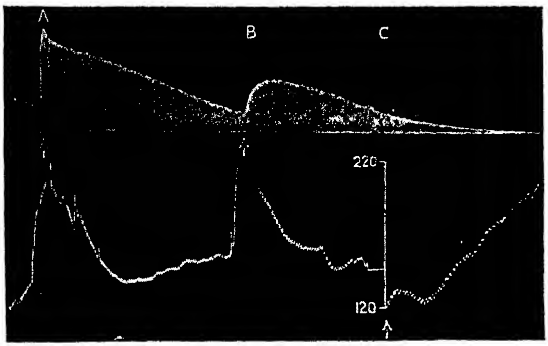


FIG. 317.—Anti-Curare Action of Adrenaline. (Wilson and Wright, *Quart. J. exp. Physiol.*, 1936, *26*.)

Above: isometric contractions of gastrocnemius in response to motor nerve stimulation; below: blood pressure. The response of the muscle has been greatly reduced by previous injection of curare, the effect of which was becoming progressively more intense. Inject (A) 60 mg KCl (cf. Fig. 314); (B) 0·02 mg adrenaline: note slow initial rise of muscle tension in response to stimulation; at C cut both vagi in the neck. Note that the three procedures produced roughly equivalent pressor effects; the pressor effect of vagal section was, however, not accompanied by any improvement in the muscle response; on the contrary, the responses decreased as the effect of the curare reasserted itself.

of skeletal muscle tone which depends on a proprioceptive reflex arc involving the *somatic nervous system alone* (p. 582).

(iii) Sympathetic stimulation *antagonizes the onset of fatigue* in skeletal muscle; its mode of action is unknown. Orbeli recorded the contraction of frog's muscle in response to tetanic stimulation of the motor nerve repeated at short intervals. Fatigue gradually set in, as shown by diminution in the height of the contraction. The sympathetic supply of the muscle was now stimulated (the motor stimulation was continued throughout the experiment); after a considerable latent period the size of the contractions became greater, and this effect persisted for some time after sympathetic stimulation was discontinued (Fig. 316). Stimulation of the sympathetic *alone* produces no recognizable effects on resting muscle.

(iv) There is evidence that *adrenaline* (which is the sympathetic transmitter (p. 508)) produces similar effects; its injection may also increase the excitability and contractility of *unfatigued* muscle above normal limits. Orbeli's results (*supra*) may be due to the sympathetic fibres to the skeletal

blood vessels liberating adrenaline, some of which diffuses out to affect the neighbouring muscle fibres.

A more striking effect of adrenaline in muscle is its *anti-curare* action which develops more slowly and is not so marked as that of K^+ ions (Fig. 317); the response is not due to the accompanying changes in the circulation. Adrenaline (and related substances like ephedrine) increase muscular power in myasthenia gravis (p. 519).

TRANSMISSION OF NERVOUS IMPULSE FROM NEURONE TO NEURONE (SYNAPTIC TRANSMISSION). TRANSMISSION IN AUTONOMIC GANGLIA.[1]

In the central nervous system and in autonomic ganglia, impulses pass from one neurone to another; conduction is always unidirectional, *i.e.* the impulse always passes from the terminals of the axon of neurone 1, to the dendrites and cell body (*soma*) of neurone 2. The "junctional" region between the terminals of neurone 1 and the surface membrane of neurone 2, is called the synapse.

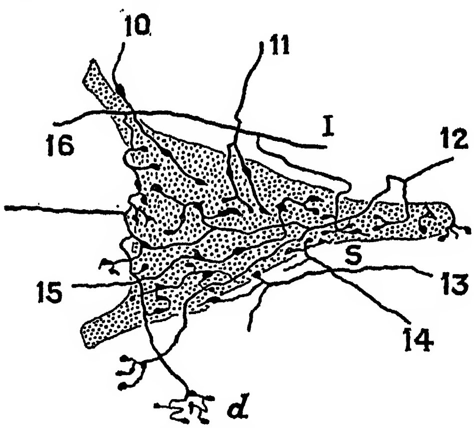


FIG. 318.—Structure of Synapse. (Lorente de No, *J. Neurophysiol.*, 1938, *1*, 195.)

The drawing shows afferent fibres ending on the surface membrane of a cell in the spinal cord.

Note: (i) the swollen synaptic terminals; (ii) the terminals of each afferent filament come in contact with a restricted part of the cell membrane.

The histology of the synapse has been most carefully studied in the grey matter of the spinal cord. The afferent nerve fibres which impinge on the surface membrane of (for example) a ventral horn cell, end in small varicosities called *synaptic terminals* (*boutons, pieds terminaux*). There is intimate contact but no physical continuity at the synapse: two membranes intervene between the two neurones, the membrane of the cell body and the membrane of the impinging nerve filaments. One ventral horn cell receives synaptic terminals from many hundreds of afferent fibres; the cell is thus a *convergence point.* Fig. 318 shows that the synaptic terminals of any *one* afferent fibre make contact with a very restricted region of the surface membrane, and do not cover this area completely; the synaptic terminals of one afferent fill up the gaps on the membrane left by the terminals of another afferent. Any *continuous patch* of the membrane of the cell body is covered by the terminals from several afferent fibres. The same kind of arrangement probably exists in the synapses of the autonomic ganglia. It should

[1] Feldberg and Gaddum, *J. Physiol.*, 1934, *81*, 305. Brown, *Physiol. Rev.*, 1937, *17*, 485. Eccles, *Ergeb. Physiol.*, 1936, *38*, 339. Bronk, *J. Neurophysiol.*, 1939, *2*, 380 Eccles, *J. Physiol.*, 1943, *101*, 465.

be noted that by contrast with synapses, the motor end plate contains only one incoming fibre and its terminals.

Transmission of the impulse from neurone to neurone ("neuro-neuronal" transmission) depends on the processes occuring at the synapse. Both the anatomical arrangements and the processes involved are simpler in autonomic ganglia than in the central nervous system. Synaptic transmission in autonomic ganglia will therefore be considered first.

Transmission in Autonomic Ganglia. Evidence for Cholinergic Transmission.—(i) When an autonomic (usually sympathetic) ganglion is perfused with eserinized saline (to preserve the acetylcholine), stimulation of the preganglionic fibres causes the release of acetylcholine, which can be

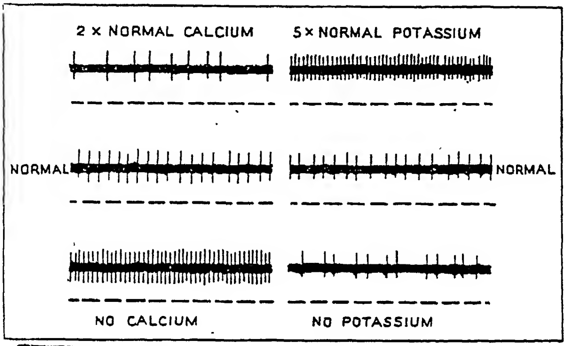


FIG. 319.—Acetylcholine stimulates Autonomic Ganglion. Effect of calcium and potassium ions on the response. (Bronk, *J. Neurophysiol.*, 1939, *2*, 293.)

The vertical lines in the tracings are action potentials (*i.e.* impulses) recorded from a *single* postganglionic fibre and thus representing the discharge of a *single* ganglion cell.

The ganglion was perfused with Ringer's solution containing 40 μg. of acetylcholine/cc.

Middle record (normal): the solution causes the ganglion cells to discharge. The Ca^{++} and K^{+} content of the perfusing fluid was modified in the upper and lower records.

Left-hand column: High Ca^{++} decreases and low Ca^{++} increases the discharge rate.

Right-hand column: high K^{+} increases and low K^{+} decreases the discharge rate.

Interrupted horizontal lines: time in 0·5 sec.

demonstrated in the fluid coming away from the ganglion by the methods described on p. 720.

(ii) Direct application of acetylcholine to ganglion cells causes them to discharge (Fig. 319); so does intravenous injection of acetylcholine as is best shown in the atropinized animal (p. 719). Atropine does not "block" transmission in ganglia, though it does "block" transmission at cholinergic postganglionic terminals. The frequency of the discharge of the ganglionic cells is increased on increasing the acetylcholine concentration employed.

(iii) K^{+} *ions* also cause the ganglion cells to discharge (they also stimulate motor end plates (p. 519)). A concentration of K^{+} ions in the perfusing fluid, which in itself is insufficient to stimulate the ganglion, enhances its response both to preganglionic stimulation and to applied acetylcholine. Fig. 319 shows that with a constant concentration of acetylcholine, the frequency of

discharge of the ganglion cell is increased by doubling the K^+ ion concentration, and is decreased by the absence of K^+ ions from the perfusion fluid. It has, therefore been suggested that (as in muscle), acetylcholine releases K^+ ions which act directly on the ganglion cell membrane. *Ca^{++} ions* have the reverse effect of K^+ ions, *i.e.* they depress transmission in ganglia (Fig. 319).

Synaptic Potentials in Ganglia.—When an impulse in a preganglionic fibre reaches a ganglion cell, it sets up a localized *synaptic potential* closely resembling in its properties the end plate potential in muscle. When the synaptic potential attains a critical level, it alters the state of the cell membrane sufficiently to generate a propagated spike potential, *i.e.* a nervous

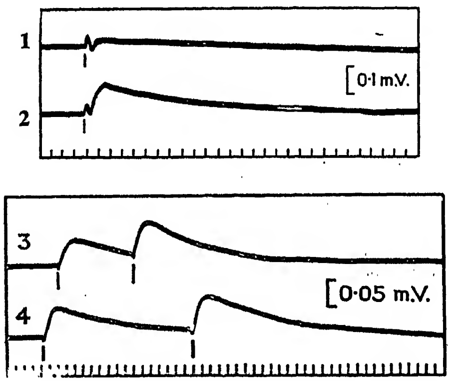


FIG. 320.—Synaptic Potential in Sympathetic Ganglion. (Eccles, *J. Physiol.*, 1943, *101*, 470, 472.)

Experiment on the stellate ganglion. The stimulus is applied to the preganglionic nerve.
Recording electrodes: the proximal is on the ganglion; the distal is on the postganglionic fibres.
The small vertical line represents the application of the stimulus. In 1 and 2 the point of stimulation is also shown by a minute deflection, the "stimulus artefact."
1. Ganglion deeply curarized. Note the small slowly rising localized (monophasic) negative synaptic potential.
2. Lighter curarization. The synaptic potential rises more abruptly and to a greater height.
3, 4. Lightly curarized ganglion. Two stimuli are applied. Summation of synaptic potentials results from the two successive preganglionic volleys of impulses.
In 4 the second stimulus is applied after a longer interval than in 3.
Time scale, 50 m.sec.

impulse, in the postganglionic fibre. It is presumed that, as in muscle, the sequence of events is as follows: nervous impulse in preganglionic fibre; release of acetylcholine at synaptic terminals; (possibly) release of K^+ ions; local (non-propagated) synaptic potential; spike potential starting in the membrane of the ganglion cell and travelling down the postganglionic fibre.

If a suitable concentration of *curare* is applied to a ganglion, a single volley of impulses, travelling along the preganglionic fibres, still releases acetylcholine which sets up a local synaptic potential only (Fig. 320, 1, 2); this potential is too small to stimulate the ganglion cells and so no discharge occurs along the postganglionic fibres. *Repetitive* stimulation of the preganglionic fibres produces summation of the synaptic potential which thus progressively increases in magnitude (Fig. 320, 3, 4). When it attains a critical size it stimulates the ganglion cells to discharge.

The ganglion cell is the cell body of a nerve fibre of the C type; as in C fibres the spike potential of the ganglion cell is followed first by a negative and then by a positive after-potential (cf. Fig. 293, C). During the negative after-potential the cell is more excitable and during the positive after-potential it is less excitable, to stimulation.

Properties of Ganglionic Transmission.—SPATIAL SUMMATION.—Any

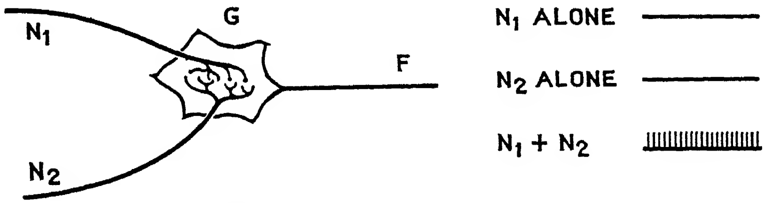


FIG. 321.—Diagram of Spatial Summation.

G = a ganglion cell; F = its postganglionic fibre.
N_1, N_2 = preganglionic fibres ending on G.
Right-hand side: record of impulses in single postganglionic fibre.
Stimulation of N_1 *or* N_2 alone: no response of ganglion cell.
Simultaneous stimulation of N_1 plus N_2: ganglion cell discharges.

single preganglionic fibre ends on several ganglion cells; it follows therefore that any single ganglion cell is a *convergence point* at which several preganglionic fibres end (Fig. 321). It is probable that before a ganglion cell can discharge, a synaptic potential of adequate *magnitude* must develop in a sufficient *area* of the surface membrane. Thus a stimulus applied to a preganglionic fibre N_1 or N_2 (Fig. 321) may produce a synaptic potential which

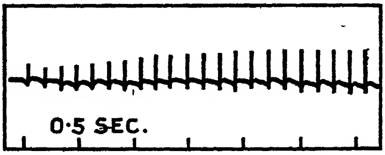


FIG. 322.—Temporal Summation in Ganglion. (Bronk, *Amer. J. Physiol.*, 1938, *122*, 8.)

Non-perfused sympathetic ganglion repetitively stimulated through its preganglionic nerve. The action potentials are recorded from the *postganglionic* nerve.
The height of the action potentials recorded is a measure of the number of postganglionic fibres in action and, therefore, of the number of ganglionic cells discharging.
Lower record: Time in 0·5 sec.
Note how with repeated stimuli the action potentials recorded from the postganglionic nerves increase in magnitude indicating that a larger number of ganglion cells is discharging.

is too small or too restricted in area to cause the ganglion cell to discharge; such a level of excitation of the cell is called *subliminal*. But if N_1 and N_2 are stimulated *simultaneously* their individual synaptic potentials are *summated*; the greater size and area of the resulting synaptic potential excites the cell and a discharge occurs. Such a process is called *spatial* summation because it involves activation of a larger area of the surface membrane (cf. p. 533).

TEMPORAL SUMMATION.—It is known that the chemical transmitter and

the synaptic potential, to which it gives rise, persist for some time. In the curarized ganglion it was shown that repetitive stimulation may build up the synaptic potential from subliminal to threshold level. A similar process possibly occurs in non-curarized ganglia (temporal summation). Fig. 322 shows that repetitive stimulation of a preganglionic nerve gradually produces an increase in the amplitude of the postganglionic spike potential, indicating that more ganglion cells are being (progressively) stimulated, *i.e.* cells which were initially stimulated subliminally, finally respond to a succession of impulses when these have built up the synaptic potential to threshold level (cf. p. 534).

After-Discharge.—Under some circumstances, on cessation of preganglionic stimulation, the ganglion cells *continue to discharge* (Fig. 323).

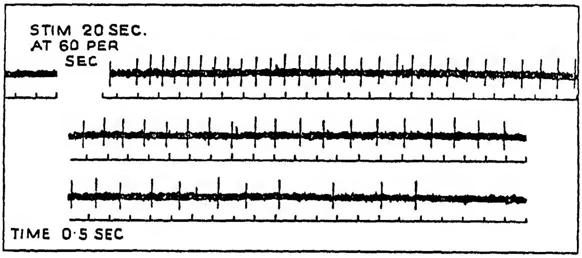


Fig. 323.—After-discharge of Ganglion Cells. (Bronk, *J. Neurophysiol.*, 1939, *2*, 387.)

Record of action potentials in postganglionic fibres, representing result of discharge of ganglion cells. During the period indicated, the preganglionic fibres were stimulated for 20 sec. at 60 per sec. On cessation of stimulation the ganglion continued to discharge (at decreasing frequency) for 27 sec. (Read record from left to right and from above downwards.) Time in 0·5 sec.

This "after-discharge" may be due either to the persistence of the transmitter which had been released during the period of stimulation or to continued release of the transmitter after stimulation had ended (cf. p. 539).

So far both temporal summation and after-discharge in ganglia have been demonstrated only under unphysiological conditions.

REFLEX ACTION. SYNAPTIC TRANSMISSION IN CENTRAL NERVOUS SYSTEM [1]

Reflex Action and the Reflex Arc.—A reflex act is usually defined as the response resulting from the passage of nervous impulses through a reflex arc. This begs the question to some extent, because the reflex arc itself must next be described. A reflex arc consists of an appropriately arranged chain of neurones, with two—*afferent* (receptor) and *efferent* (effector,

[1] Sherrington, *Integrative Action of Nervous System*, new edn., Cambridge, 1947. Creed, Denny-Brown, Eccles, Liddell, and Sherrington, *Reflex Activity of Spinal Cord*, Oxford, 1932. *Selected Writings of Sir Chas. Sherrington*, ed. Denny-Brown, London, 1939. Lorente de No, *J. Neurophysiol.*, 1939, *2*, 402. Lloyd, *Physiol. Rev.*, 1944, *24*, 1. Eccles, *J. Neurophysiol.*, 1947, *10*, 251.

excitor)—as the minimum. The afferent neurone leads from the receptor (sensory) organ (cutaneous, muscular, or special sense) via a dorsal nerve root or its cranial equivalent into the central nervous system; the nutrient cell is in the dorsal root ganglion (or cranial equivalent). The efferent neurone supplying skeletal muscle has its cell body in the central nervous system (ventral horn cell or motor cranial nucleus), and the efferent fibre passes out in the ventral nerve root (or motor cranial nerve) to the muscle (effector organ). (In the case of viscera the efferent side of the arc has a more complex arrangement, cf. p. 704.) In only a few spinal reflexes (one of these is the stretch reflex) is the minimal number of two neurones, involved; generally one, several, or many connecting or *internuncial neurones* (*interneurones*) intervene between the afferent and efferent neurones. The interneurones may be situated in the same segment of the spinal cord, or the afferent fibre may pass up or down in the spinal cord for varying distances before relaying; ultimately the interneurones establish connections with the effector neurones (cf. p. 529).

The effector neurones (*motor neurones*), called clinically the *lower motor neurones, e.g.* ventral horn cells and fibres, constitute the *final common path* or sole available route to the skeletal muscles. They are the sole path which all impulses, no matter whence they come, must travel, if they are to act on the muscle fibres to which they lead. Afferent impulses originating at different distant points and passing through paths widely separate in the brain, converge to the same motor mechanism—the final common path, and act *harmoniously* upon it.

The definition of reflex action given above is essentially anatomical and is too general in character because it could easily be made to include almost every form of neural activity; thus the most complex and deliberate voluntary act is, on ultimate analysis, the response to impulses which have at some preceding time entered the central nervous system. Certain other criteria are therefore often employed. The response must be *inborn* and be present in all members of the species. These additional considerations exclude the so-called *conditioned reflexes* (p. 675) or habits and other motor acts which are acquired after birth as the result of a process of conditioning, training, or learning. The moment of birth is an arbitrary limiting time, because the central nervous system is not anatomically completed in the higher animals till later, *e.g.* some of the long tracts have not yet acquired medullary sheaths along their entire course. A reflex response is said to occur "without the necessary intervention of consciousness," and to be "involuntary." Unfortunately, it is very difficult to define a voluntary act (cf. p. 646). Certain acts which are generally thought of as voluntary, *e.g.* walking, may be carried out successfully with a minimum of attention and with the mind devoted to other matters. Stress is also laid on the *level* in the nervous system through which the arc passes, *i.e.* the so-called "centre" for the reflex. If the "centre" is in the spinal cord, brain stem, or basal ganglia, the reflex nature of the response is not questioned, but if it is in the cerebral cortex doubts may arise.

Reflexes are often described as machine-like and inevitable responses to afferent stimuli. This description is true when only the general character of the reaction is considered, but when examined in detail the reflex response usually shows a considerable measure of variability and is modified by

environmental changes, and especially by the "previous history" of the reflex centre. *Complexity* of response is not a criterion; for some reflexes are very elaborate and may involve the greater part of the body. Generally speaking, reflex responses become more *variable* in pattern as the centre lies at *higher* levels in the central nervous system.

Transmission in the Central Nervous System.—Two fundamental differences exist between autonomic ganglia and the central nervous system.

(1) All preganglionic fibres of the autonomic system are excitatory to ganglion cells; *i.e.* when preganglionic impulses reach the synaptic terminals, stimulation of ganglion cells occurs. Afferent fibres in the central nervous system on the other hand may be either *excitatory* or *inhibitory*. An excitatory afferent, when stimulated, causes a discharge from central ventral horn cells and consequently, contraction of the corresponding muscle fibres; when an inhibitory afferent is stimulated it diminishes or arrests the discharge of certain ventral horn cells and the corresponding muscle fibres consequently relax. Two types of central transmission therefore occur, *excitatory transmission* and *inhibitory transmission* (Fig. 324).

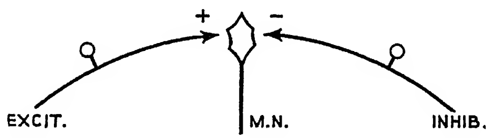


FIG. 324.—Antagonism of Central Excitation and Central Inhibition.

M.N. = Motor Neurone.
Excit., Inhib. = afferent excitatory and afferent inhibitory fibre.

Although central transmission has been intensively studied the exact nature of the process is still obscure. As central excitation and central inhibition can mutually annul or neutralize one another they may be due to processes which are opposite in sign.

(i) *Central Excitation.*—Arguing from the analogy of the end plate potential and the ganglionic synaptic potential, it will be assumed here that central excitation is ultimately due to the development of a synaptic potential on the affected cell. This potential is conceived as a localized *negative* electrotonus (*catelectrotonus*); when it involves a sufficient area of the surface membrane of the cell and attains an adequate magnitude it causes excitation (discharge) of the cell.

(ii) *Central Inhibition.*—Conversely central inhibition is here attributed ultimately to the development of a localized *positive* electrotonus (*anelectrotonus*) on the surface membrane of the cell. The anelectrotonus annuls the catelectrotonus and so prevents excitation (discharge) of the cell.

Neither the occurrence, nor the identity of chemical transmitters in the central nervous system has been proved beyond doubt. There is some evidence suggesting that acetylcholine might act as a central synaptic transmitter; it is conceivable that acetylcholine might, in certain circumstances,

act as an excitatory transmitter (E) and in other circumstances, as an inhibitory transmitter (I) (p. 530).[1]

(2) INTERNUNCIAL NEURONES (INTERNEURONES).—In autonomic ganglia, the preganglionic fibres end directly on the ganglion cells. In the central

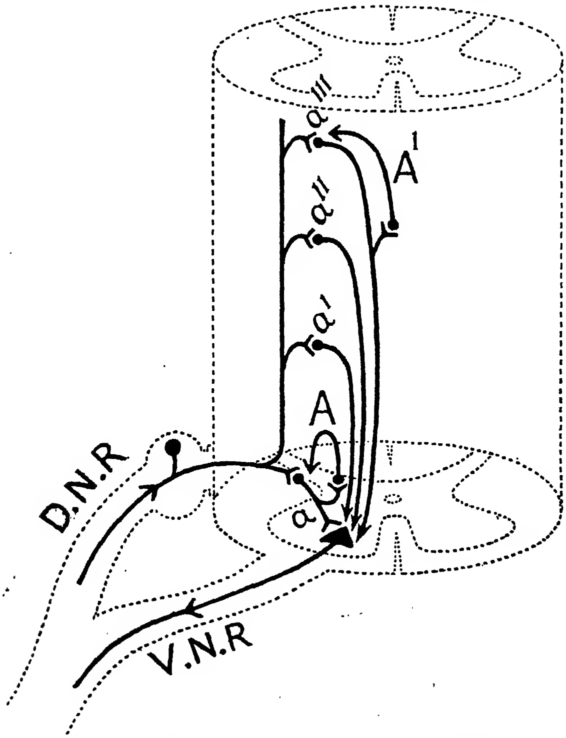


FIG. 325.—Rôle of Internuncial Neurones in Reflex Action.

Diagrammatic section of spinal cord. D.N.R.=dorsal (posterior) nerve root; V.N.R.=ventral (anterior) nerve root; *a*, *a′*, *a″*, *a‴*, A, A′=internuncial neurones.
The shortest reflex arc is via D.N.R., *a*, V.N.R. Longer reflex arcs involving delay paths are shown via *a′*, *a″*, or *a‴*. A, A′, are "reverberators." Note how an impulse passing from D.N.R. to V.N.R. along *a* may branch to excite A, which in turn re-excites *a* ;similarly an impulse along *a‴* may branch to excite A′, which in turn re-excites *a‴*.

nervous system, as explained on p. 527, one or a few or many internuncial neurones generally intervene between the primary afferent neurones and the motor neurones. Afferent impulses entering the spinal cord, may thus flow

[1] Sherrington has pointed out that the existence of " a surface of separation " at the synapse makes possible a wide variety of local changes that might be involved in excitatory or inhibitory transmission. " Such a surface might restrain diffusion, bank up osmotic pressure, restrict the movement of ions, accumulate electric changes, support a double electric layer, alter in shape and surface tension with changes in difference of potential, alter in difference of potential with changes in surface tension or in shape, or intervene as a membrane between dilute solutions of electrolytes of different concentration or colloidal suspensions with different signs of charge " (*Integrative Action of Nervous System*, new edn., 1947). Sherrington thus envisages a much wider range of possible processes concerned with central transmission than has been considered in the text.

through many channels before reaching the ventral horn cells. The inter-neurones may play a considerable part in the processes of central summation and after-discharge. Consider, for example, the arrangement of the inter-nuncials in the reflex arc illustrated in Fig. 325. A single stimulus applied to an appropriate afferent nerve sets up a single afferent impulse which enters the spinal cord in the dorsal nerve root. At the branching point of this root the impulse takes the shortest route to the internuncial *a* and then to the ventral horn cell, using a three-neurone reflex arc. The impulse will also travel along the ascending branch in the spinal cord and ultimately reach the ventral horn cell after having travelled along progressively longer reflex arcs via internuncials *a′*, *a″*, and *a‴*. The motor cell will thus be bombarded by *four* impulses reaching it in close succession although only a *single* afferent impulse was set up in the dorsal root. Consider further the internuncials A and A′. An impulse travelling along the internuncial *a* will not only stimulate the ventral horn cell but also the cell of A; the impulse along the fibre of A excites the cell of *a*. An excitatory cycle is thus established by means of which *a* stimulates A, and A in its turn stimulates *a*. Each time the cycle is completed the motor neurone is stimulated (cf. *circus movement* in heart, p. 291). In this way (as the nerve impulses travel without decrement) the motor neurone might be subjected to a nervous bombardment for an endless period of time unless some inhibitory agency supervened. The neurone A, and the similarly arranged internuncial A′ may conveniently be called *reverberators*. Reverberation or sustained activity so produced in the nerve centres leads to *after-discharge* of the motor neurones and offers opportunities for facilitation or central summation as explained on p. 535 and p. 539. It should be remembered that the conventional 2- or 3-neurone reflex arc is rarely used in the body and that normally arcs of greater complexity are employed.

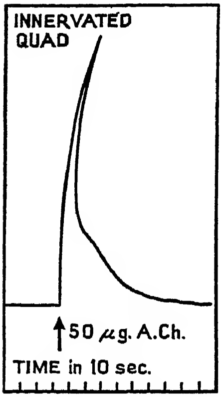


FIG. 326.—Stimulating (excitor) action of Acetylcholine on Spinal Cord. (Calma and Wright, *J. Physiol.*, 1944, *103*, 95).

Decerebrate cat. Tension of quadriceps of innervated leg. At signal inject 50 μg. of acetylcholine into central end of subclavian artery so as to reach spinal cord. Note rapid vigorous contraction of innervated quadriceps which is due to acetylcholine setting up a nervous discharge from the ventral horn cells of the spinal cord.

Pharmacology of Central Transmission.[1]—(1) ACTION OF ACETYLCHOLINE ON SPINAL CORD.—The effect on the spinal cord varies with the exact experimental conditions, and may be *excitatory* or *inhibitory*. In the decerebrate cat, intra-arterial injection of acetylcholine stimulates ventral horn cells and causes a discharge of nerve impulses and contraction of the corresponding muscle fibres (Fig. 326). In the cat under chloralose anæsthesia, acetylcholine generally inhibits reflex spinal activity; the knee-jerk, for example, is diminished or abolished (Fig. 327). This effect is not due to the

[1] Schweitzer and Wright, *J. Physiol.*, 1937, *89*, 165, 384; *90*, 310; 1938, *92*, 422. Schweitzer, Stedman, and Wright, *J. Physiol.*, 1939, *96*, 302. Calma and Wright, *J Physiol.*, 1944, *103*, 93. Calma, *J. Physiol.*, 1949, *108*, 282.

associated fall of blood pressure or to changes in the muscles themselves, but to a direct inhibitory action on the spinal cord. The central action of acetylcholine is partially annulled by atropine.

(2) ACTION OF ANTICHOLINESTERASES OF URETHANE TYPE ON SPINAL CORD.—The anticholinesterases modify reflexes as a result of a direct action on the nerve elements in the spinal cord. The *tertiary* anticholinesterases (*e.g. eserine sulphate*) usually stimulate the spinal cord (Fig. 328) enhance the reflexes, and may give rise (in larger doses) to generalized and powerful convulsions. The *quaternary* anticholinesterases, *e.g. prostigmine*, however, under appropriate experimental conditions may have just the reverse effect and depress reflexes by a direct action on the spinal cord. Prostigmine, for example, *injected intrathecally in man* abolishes muscle tone and reflexes and diminishes the strength of voluntary movement by such a direct central action.

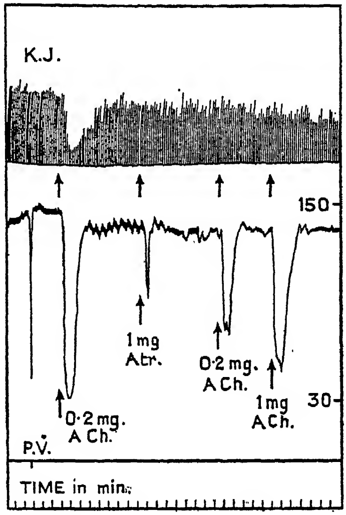


FIG. 327.—Action on Spinal Cord of Acetylcholine and Atropine. Central Inhibitory Action of Acetylcholine is partially annulled by Atropine. (Schweitzer and Wright, *J. Physiol.*, 1937, *89*, 183.)

Records from above downwards are knee jerk, blood pressure, base line, time in minutes. At first signal PV on base line stimulate peripheral end of vagus; note usual sharp fall of blood pressure.
At first arrow inject 0·2 mg. of acetylcholine; note *inhibition of knee-jerk*. At second arrow inject 1 mg. of atropine.
At third and fourth arrows, again inject acetylcholine; there is no depression with 0·2 mg. and a slight gradual depression of the knee-jerk with 1 mg.

(iii) *Differences between Convulsant and Depressant Compounds.*—These differences in the central action of the two groups of anticholinesterases occur although the drugs have identical actions on *other* tissues. Thus the anticholinesterases D.C. hordenine hydrochloride (tertiary) and the corresponding methiodide (quaternary) have similar actions (both quantitative and qualitative) on cholinesterase *in vitro*, on skeletal muscle, blood pressure, and other functions; but the former is a central convulsant and the latter a central depressant (Fig. 329). A possible explanation of these differences in their action on the spinal cord is as follows: The *quaternary* ammonium anticholinesterases and their derivatives are lipid-insoluble and so probably cannot *penetrate the lipid envelopes* of the nerve cells; their action is consequently probably limited to the external surface of these cells. The *tertiary* ammonium compounds on the other hand give rise to free base which, being lipid-soluble, may perhaps dissolve in the lipid material of the cell surface and enter the interior of the cell. It is suggested that the anticholinesterase which penetrates into the cell acts as a convulsant, while a member of the group which is unable so to penetrate is a central depressant. The central action of both the convulsant and depressant anticholinesterases may thus

be due to an increase in the *acetylcholine concentration at specific regions of the grey matter*. The difference between the two groups may depend on the site of their activity in relation to the surface membranes of the nerve cell. These results suggest the possibility that acetylcholine could act both as a central excitatory and central inhibitory agent within the spinal cord, and that the level of activity of ventral horn cells may perhaps depend on the relative concentrations of acetylcholine within and without these cells.

The other groups of anticholinesterases, *i.e.* diisofluorophosphonate, and the simpler polyphosphates, are all central excitants (like eserine sulphate).

It has been suggested (partly on the basis of the evidence quoted above) that acetylcholine may be concerned in central transmission; other transmitters may also be involved.[1] It is particularly interesting to note that adrenergic fibres have been demonstrated in the central nervous system. Their rôle is unknown.

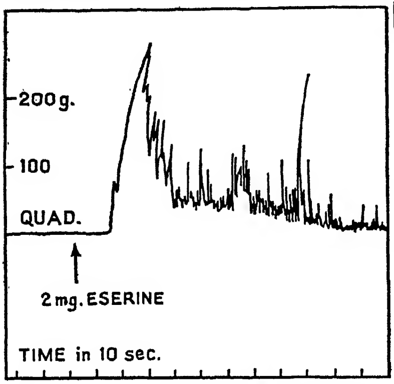


FIG. 328.—Stimulating (Excitor) Action of Eserine on Spinal Cord. (Calma and Wright, *J. Physiol.*, 1944, *103*, 101.)

Decerebrate Cat. Record of Tension of Innervated Quadriceps.
Animal fully atropinized.
At arrow, inject intravenously 2 mg. of eserine.
The spinal ventral horn cells are stimulated and discharge to the quadriceps which contracts first tonically and then clonically.

General Features of Reflex Excitation.—(1) SYNAPTIC DELAY.—This is the time taken in transmitting the impulse through one synapse. It has been measured thus (Fig. 330): a stimulating electrode is introduced into the grey matter of the 4th cranial nucleus which contains both interneurones and motor neurones. A recording electrode (R) is placed on the 4th cranial nerve to record the arrival of impulses set up in the motor neurones (M). If a *weak* stimulus is employed, it stimulates the interneurones (I) only, and the impulse must first *pass through one synapse* to excite the motor neurones (M); if a *strong* stimulus is employed, the motor neurones are also stimulated *directly* by the stimulating current and a motor impulse is discharged sooner. The time interval between the two arrows (Direct M, M via I) is the time taken for the impulse to travel along a very short interneurone (I) and then through the synapse to impinge on the motor neurones (M). The delay recorded varies with the degree of excitability of the motor neurone and is 0·5–1·0 m.sec. As some of the recorded delay is due to conduction along the slowly conducting very fine terminal fibres of the interneurone the true synaptic delay must be even more brief (*e.g.* 0·1–0·2 m.sec.).

(2) REFLEX DELAY.—The *total reflex time* is the interval between the application of the stimulus to the peripheral afferent nerve and the onset of contraction in the muscle. Part of the delay is due to the time taken by the volley of impulses to travel in the peripheral nerves to and from the centre, and the latent period of the muscle. This *peripheral time* can easily be measured; when deducted from the total reflex time it leaves the *central*

[1] For critical discussion see Feldberg, *Brit. med. Bull.*, 1950, *6*, 312.

reflex time or *central delay*, *i.e.* the time taken by the transmission processes in the central nervous system.

The length of the central delay is probably an index of the number of neurones employed in the reflex arc, *i.e.* the number of synapses that are successively traversed. In Fig. 331, A, a muscle afferent was stimulated to produce the stretch reflex (p. 588); in Fig. 331, B, a cutaneous afferent was stimulated to produce the flexor reflex (p. 692). The potentials were recorded in each case from the ventral nerve root; the total latency of the stretch reflex (from time of stimulation to the appearance of the ventral root spike potential) was 2·5 m.sec. while that of the flexor reflex was 5·5 m.sec. although the peripheral delay was approximately the same in the two cases. When the central delay in these experiments is calculated it can be concluded that the stretch reflex probably employs a 2-neurone arc involving transmission at one synapse only; the flexor reflex on the other hand employs a multi-neurone arc, involving, perhaps, 4 or 5 successive synapses.

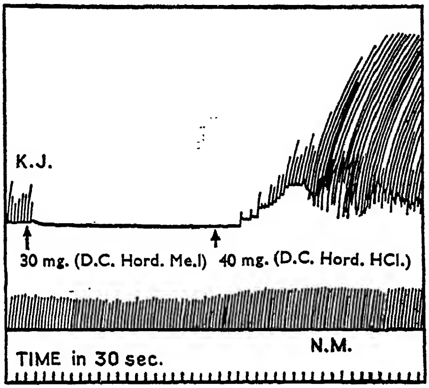


FIG. 329.—Contrasted Central action of Tertiary and Quaternary Anticholinesterases on Activity of Spinal Cord. (Schweitzer and Wright, *J. Physiol*, 1938, *92*, 431).

Tertiary compound = dimethyl carbamic ester of hordenine *hydrochloride*. (D.C. Hord. HCl.)
Quaternary compound = dimethyl carbamic ester of hordenine *methiodide*. (D.C. Hord. Me.I.)
Cat. Chloralose. Records from above downwards are: K.J. = knee-jerk (right side); N.M. = contractions gastrocnemius (left side), stimulated through its motor nerve; time in 30 secs.
First arrow: inject 30 mg of D.C. Hord. Me.I; the knee-jerk is abolished from a depressant action on spinal cord. Second arrow: inject 40 mg of D.C. Hord. HCl. Progressive increase in muscle tone, return and later great increase in knee-jerk, development of violent convulsive movements. These changes are due to stimulation of spinal cord.

(3) CHANGES IN THE NERVE CELL AS A RESULT OF ACTIVITY.—The nerve impulse is conducted over the cell in the same way as along the nerve fibre. Activity in the cell is accompanied and signalled by a spike potential; there is no following phase of heightened excitability; throughout the recovery period the cell shows depressed excitability (cf. positive after-potential in nerve (p. 489)).

(4) CENTRAL SUMMATION.—As in autonomic ganglia, summation might be (*a*) *spatial*; (*b*) *temporal*. It is assumed in this discussion that the synaptic potential is directly responsible for exciting the cell; if it attains threshold value (in magnitude and extent of cell membrane involved), it "fires" the cell (causes it to discharge); if it is subliminal, it fails to fire the cell. But a subliminal synaptic potential may summate with another subliminal synaptic potential which has been set up simultaneously in an adjacent part of the cell membrane, and thus reach threshold level; this process is *spatial summation*. Further, as the synaptic potential endures for some time, (although it is a very brief period), a sequence of impulses might conceivably build up subliminal potentials, stepwise, to threshold value—this would be *temporal summation*; it is extremely doubtful whether true temporal summation ever occurs in the central nervous system.

(i) *Spatial Summation.* This phenomenon is well demonstrated in Fig. 332, in which the scratch reflex was studied. The reflex is so named because it is elicited naturally by scratching the skin on the back (dog); the rhythmic hind limb movement produced tends to remove the irritant. When *either* of the two skin points α or β is stimulated alone, the afferent impulses set up fail to elicit a reflex discharge of the motor neurones; in relation to the *reflex* therefore, the skin stimulus employed is subliminal (though to the skin nerve endings the stimulus is adequate). When α and β are stimulated *simultaneously* the motor neurones discharge and the hip flexor contracts rhythmically (scratch reflex).

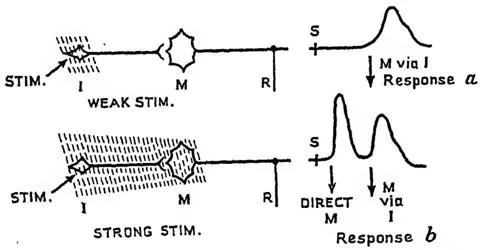


FIG. 330.—Synaptic Delay of Motor Neurones. (Modified from Lorente de No, *J. Neurophysiol.*, 1939, *2*, 402.)
Left-hand diagrams: I=interneurone. M=motor neurone (4th cranial nucleus). R=recording electrode on motor nerve fibre (4th cranial nerve).
Right-hand records: Electrical responses recorded from electrodes applied to the motor nerve (4th cranial). S=moment of stimulation).
Upper record: Apply a *weak* stimulus to excite the *interneurones* only. The impulse which is set up must therefore travel to M and pass through one synapse to stimulate the motor neurones. Note single response (M via I).
Lower record: A *stronger* stimulus is applied which simultaneously excites both the interneurones and motor neurones directly. The motor neurones are thus excited first directly (Direct M), and subsequently indirectly by the impulse set up at I (M via I). Note double response on the lower electrical record. The time interval between the first arrow (Direct M) and the second arrow (M via I) is the synaptic delay.

It is supposed that the fibres from α and β end (in part at any rate) on a common group of motor neurones. Impulses from α or β alone set up there a subliminal level of "central disturbance" or "central excitatory state"; or in terms of the hypothesis adopted here, they set up a synaptic potential which is too small in magnitude or extent to "fire" the motor neurones. When impulses from α *and* β reach the motor neurones simultaneously spatial summation occurs; the synaptic potentials may be supposed to be built up to threshold level and the motor neurones discharge.

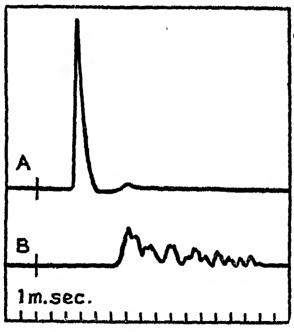


FIG. 331.—Central Delay depends on Number of Synapses in Reflex Path. (After Lloyd, *J. Neurophysiol.*, 1943, *6*, 111.)
A. Stretch Reflex: total reflex delay=2·5 m.sec.
B. Flexor Reflex: total reflex delay =5·5 m.sec. Potentials are recorded from ventral root.

(ii) *Temporal Summation* (Fig. 333). An afferent nerve is stimulated once at such a strength that although a volley of impulses is set up, no firing of motor neurones occurs; *i.e.* with respect to the reflex, the stimulus is subliminal. Repetition of the stimulus (at the same strength) to the same nerve at suitable intervals (*e.g.* 3 per sec.) and for an adequate number of times (*e.g.* 6 times), owing to central summation, finally causes firing of motor neurones and muscular contraction (flexor reflex) results.

This experiment requires critical analysis. It could be interpreted

simply in terms of temporal summation of the synaptic potential. Let us assume that the reflex arc consists of two neurones only and, therefore, involves transmission at a single synapse (Fig. 334, A). The first afferent volley sets up a subliminal synaptic potential which rises to a peak and declines (like the ganglionic potential, Fig. 320). Before this potential has disappeared the next volley arrives setting up a second subliminal synaptic potential which summates with what is left of the first. Successive volleys result in progressive building up of the synaptic potential until it finally attains threshold value and "fires" the motor neurones.

This interpretation is probably incorrect for the following reasons:

(*a*) Direct experiment (p. 532) suggests that the duration of the synaptic potential at motor neurones is less than 1 m.sec., probably 0·1–0·2 m.sec. The interval between successive stimuli in the experiment illustrated by Fig. 333 was *30 m.sec.*: simple summation of the synaptic potential could thus not have occurred.

(*b*) The flexor reflex involves several internuncials (Fig. 331) and its reflex arc resembles that shown in Fig. 325, or in simplified form in Fig 334, B.

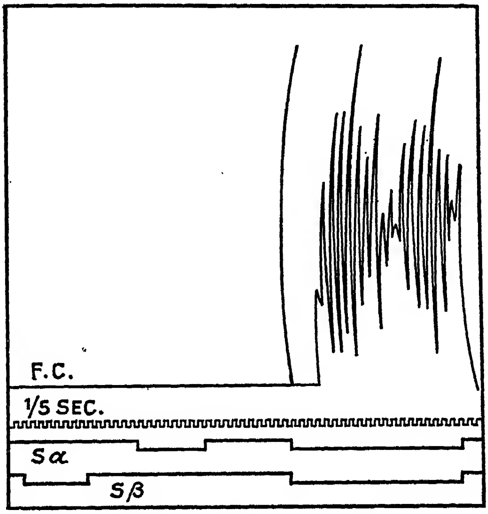


FIG. 332.—Reflex (Spatial) Summation in the Scratch Reflex. (Sherrington, *Integrative Action of Nervous System*, Cambridge, new edn., 1947.)

Records from above downwards.
FC: Record of contraction of flexor muscle of hip.
Time in 1/5 sec.
Sα, Sβ: *descent of signal line* represents period of stimulation of point α or β on the shoulder skin. (If either α or β is *strongly* stimulated it elicits rhythmic flexion of the hip, *i.e.* the scratch reflex.)
First signals on Sα, Sβ: *weak* stimulation of skin points α or β applied separately elicits no reflex response.
Second signals on Sα, Sβ: same weak stimulation of skin points α and β applied *simultaneously*. Spatial summation occurs and a powerful rhythmic scratch reflex results.

Suppose afferent L in Fig. 334, B, is stimulated repetitively. The first volley stimulates the *interneurones* as well as the motor neurones. When the second afferent volley enters the cord it finds the motor neurones in a state of subliminal excitation because of impulses reaching them at *about the same time* from the activity of the internuncials. With each successive volley one may suppose that a larger internuncial pool discharges, raising the excitatory state of the motor neurones nearer to threshold. "Firing" occurs when the excitatory state of the motor neurones has been raised in this way to such a level that the additional excitation (or synaptic potential) set up by an afferent volley *arriving at this time* brings the excitatory state up to threshold level. In other words the summation that is taking place is really *spatial*; repetition of the stimuli raises the excitatory state of the

motor neurones *not* by simple summation in time (true temporal summation) but by building up an adequate level of "background bombardment" from the internuncials. The effects of an afferent stimulus are often determined largely by this background of internuncial activity.

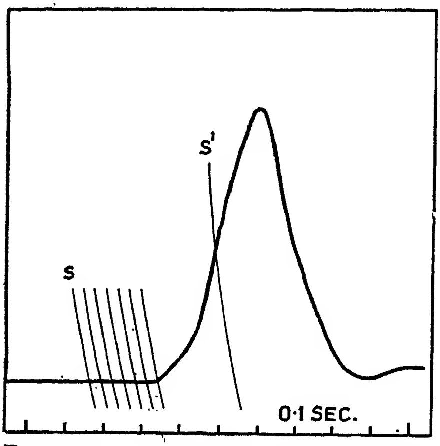


FIG. 333.—"Temporal" Summation of Reflexes. (After Sherrington, *Integrative Action of Nervous System.*)

Flexor Reflex (contraction of knee flexor muscles). Each vertical line S represents a subliminal stimulus applied to ipsilateral afferent nerve. After six such stimuli the motor neurones fire and contraction of the muscle sets in. Stimulation ended at S¹. Time (lower record) in 0·1 sec.

(5) BLOCK OR RESISTANCE.—A nervous impulse cannot pass backwards through a synapse, *i.e.* it can pass from the terminations of the axon to the dendrites of the adjacent neurone, but not in the reverse direction. In this connection the peripheral fibres of the dorsal root ganglia and their sensory cranial homologues must be regarded as equivalent to dendrites, while the central fibre is the true axon. The term *law of forward conduction* is applied to this fixity of direction of impulse-conduction in the nervous system.

(6) LOCALIZATION.—To elicit any reflex the stimulus must be applied to a particular locality; *e.g.* to elicit the flexor reflex an afferent nerve in the ipsilateral hind limb must be stimulated (Fig. 333).

(7) EFFECTS OF INCREASING STRENGTH OF STIMULATION.—(i) As the strength of afferent stimulation is progressively increased the reflex response under examination shows the following changes: the latent period becomes shorter; the rate of development of tension and the peak tension increase; after-discharge is prolonged (Fig. 335).

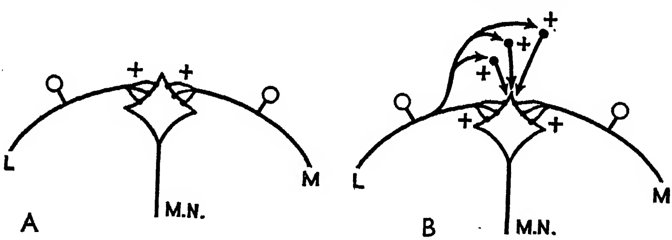


FIG. 334.—Diagram of Mechanism of Temporal Summation.

A. Stimulation of afferent nerve L or M involves one synapse only.
B. Stimulation of afferent nerve L involves interneurones.
M.N., motor nerve.

(ii) When the response of the animal as a whole is being studied, it is found that stronger afferent stimulation produces a more widespread response (*irradiation*). Thus, weak stimulation of the sole of the foot elicits only toe flexion; stronger stimulation may give rise to flexion of the whole hind limb,

i.e. a greater part of the flexor motor centre (*i.e.* the cells innervating all the flexor muscles) is activated. This is due to a larger number of afferent fibres being stimulated, thus activating more motor neurones.

(8) Fatigue.—Reflex responses show fatigue comparatively readily; the latent period becomes longer and the rise of tension smaller and more gradual. Reflex fatigue is due to some change developing in the centre: thus, when the flexor reflex can no longer be elicited, the peripheral mechanism is still active, and stimulation of the motor nerve readily elicits muscular contraction. If the blood supply of the centre is impaired, or if it is depressed by anoxia or anæsthetics, fatigue occurs sooner. The intimate nature of central fatigue is however unknown.

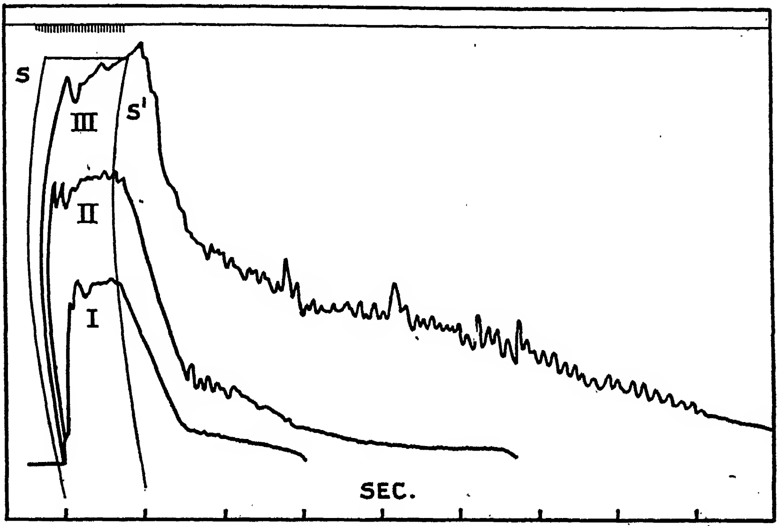


Fig. 335.—Effect of Intensity of Stimulus on Magnitude of Response and on After-discharge of Flexor Reflex. (Sherrington, *Integrative Action of Nervous System*, Cambridge, new edn., 1947.)

Flexor Reflex. Afferent nerve stimulation during the period S–S¹. The strength of the stimulus was progressively increased in the order I, II, and III. Note that the stronger stimulus produces a greater immediate response and a longer after-discharge. Time in seconds.

(9) Rebound is a phenomenon peculiar to reflex action, and has not yet been accounted for. On cessation of reflex inhibition of a muscle, its tension may rise considerably above its previous resting value (Fig. 344); this rebound contraction is independent of previous elongation of the muscle, for it may take place when the muscle is so toneless that lengthening does not occur during the period of inhibitory afferent stimulation. Rebound is attributed to an "overswing" in the centres—a state of over-excitation following on a state of inhibition; but this "explanation" puts more questions than it answers.

(10) Recruitment and After-Discharge.[1]—These features are well brought out when the *crossed extensor reflex* (*e.g.* contraction of the quadriceps

[1] Liddell and Sherrington, *Proc. roy. Soc. B.*, 1925, *97*, 488.

as a result of stimulation of an afferent nerve in the *opposite* limb) is contrasted with the motor tetanus of the same muscle [1] (Fig. 336).

(i) The motor nerve is stimulated for a few seconds at a frequency sufficiently high to produce complete tetanus. After a brief latency the tension developed by the muscle rises *sharply* to a maximum; when the stimulus is discontinued, the tension diminishes *rapidly* as the muscle relaxes (Fig. 336, B).

(ii) When the *crossed extensor reflex* is elicited there is a *longer latency*, which is characteristic of reflex activity; under continued afferent stimulation the tension in the quadriceps rises relatively *gradually* to its maximum. When stimulation is stopped the tension is *maintained* for some time and then *slowly declines* (Fig. 336, A). The explanation of these differences is as follows:

As the muscle fibres are under the same mechanical and nutritive conditions in both experiments and the same tension develops on both occasions, the same number of muscle fibres may be presumed to be contracting in

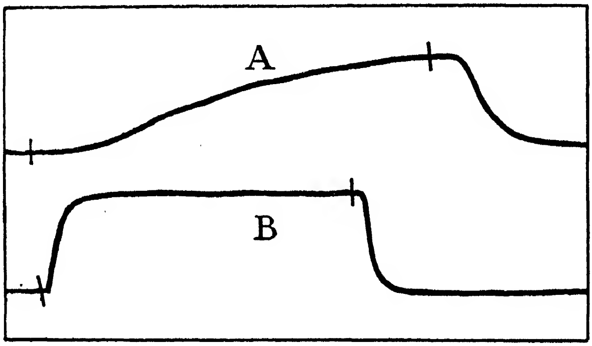


FIG. 336.—Crossed Extensor Reflex compared with Motor Tetanus. (Liddell and Sherrington, *Proc. roy. Soc. B.*, 1923.)

Quadriceps muscle. A: Crossed extensor reflex produced by contralateral afferent stimulation, frequency 38 per sec.; B: Tetanus of muscle produced by peripheral motor nerve stimulation, frequency 38 per sec. (isometric records). The vertical lines indicate the beginning and the end of stimulation.

both reactions. The rapid rise of tension in the motor tetanus indicates that all the muscle fibres involved contract practically *synchronously*; the slower development of tension in the reflex contraction suggests that first a few motor units are affected, then gradually more and more become involved, until the full quota is in action. In other words, with continued excitation of an afferent nerve, an increasing number of motor neurones is brought under excitation: this is called *excitatory recruitment*. Recruitment is due to "*pseudo*-temporal summation," *i.e.* with repetitive afferent stimulation there is a progressive increase in the "background" activity of the internuncials. This leads to an increase in the excitability of more and more motor neurones until *spatial* summation raises the local synaptic potential to threshold, causing firing (p. 535). If afferent stimulation is continued for a long time, effective summation takes place at a steadily increasing number of motor neurones. Indefinite prolongation of the afferent stimulus does not produce unlimited recruitment; a *stimulation plateau* is

[1] These features are often less well shown by the flexor reflex

reached (Fig. 336). There is thus a limit to the number of motor neurones that can be recruited.

On discontinuance of afferent stimulation the tension may remain unaltered for several seconds; this is called the *after-discharge plateau.* As it is equal in height to the stimulation plateau, *all* the motor neurones which were ultimately excited during afferent stimulation must still be discharging. When relaxation of the muscle sets in, it proceeds more gradually than in the motor tetanus. In the latter all the muscle fibres go out of action together; in the case of the reflex it must be supposed that the motor neurones stop discharging successively and as this happens the corresponding muscle fibres relax. Recruitment thus gives *inertia,* and after-discharge provides *momentum* to reflex movements, making these movements smoother in onset and termination.

Fig. 337 shows the after-discharge in the flexor reflex elicited by a *single* afferent stimulus. Action potentials (representing the results of motor neurone discharge) are present right to the end of muscular relaxation. The motor

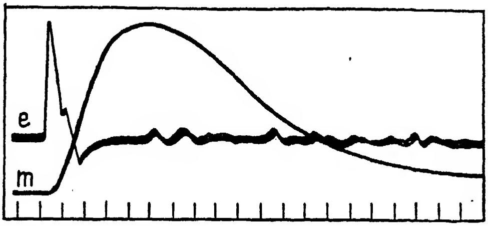


FIG. 337.—Reflex After-Discharge. (Sherrington and others, *Reflex Activity of Spinal Cord,* 1932.)

Mechanical (m) and electrical response (e) to *single* shock to afferent nerve producing reflex response (cf. with Fig. 302, A). A large number of action potentials are visible in the electrical record right up to the end of muscular activity. These represent the after-discharge of the spinal ventral horn cells. Relaxation is consequently more gradual than in a motor nerve twitch. Time (on base line) in 0·01 second.

neurones thus continue to discharge long after afferent stimulation has ceased. After-discharge is attributed to *persistent stimulation of the motor neurones from the internuncial background.* As has been repeatedly stated, besides taking the direct short route to the motor neurones, impulses may wander through long and tortuous internuncial "delay paths" in the intricate maze of the central nervous system before ultimately reaching their final objective, the motor neurones (Fig. 325). Impulses go on wandering in these paths for varying periods after afferent stimulation ceases, continue to bombard the motor neurones, and so maintain the after-discharge. As already explained (p. 530), there is no reason on anatomical grounds why such central reverberation should not continue for a long time.

It is easier reflexly to inhibit the after-discharge than the stimulation plateau. This is well seen in Fig. 338, where stimulation of an inhibitory afferent produces only trifling relaxation during the stimulation plateau but causes complete and rapid relaxation during the after-discharge. This can be readily understood because the after-discharge is due solely to the impulses along the longer delay paths (the "stragglers"), while the stimulation plateau is due to impulses arriving simultaneously by all available routes.

It follows from what has been said that prolongation of afferent stimulation increases (within limits) the peak tension and the duration of the after-discharge.

(11) Occlusion.—When two afferent excitatory nerves (*a* and *b*)—each of which can evoke the flexor reflex—are simultaneously stimulated, it is sometimes found that the tension developed by the flexor muscle under observation is *less* than the sum of the tension produced by each afferent stimulated separately; thus if *a* produces in the muscle a tension of value 9 (in arbitrary units), and *b* also a tension of 9, stimulation of *a* and *b* together may only yield a tension of 12 (instead of the expected 18). This phenomenon is referred to as occlusion; it is due to the fact that some of the spinal motor neurones (in this case producing 6 units of tension in the muscle) are common to both *a* and *b*. As these motor neurones are maximally excited when *a* or *b* is stimulated *separately*, they naturally give no greater response when *a* and *b* are

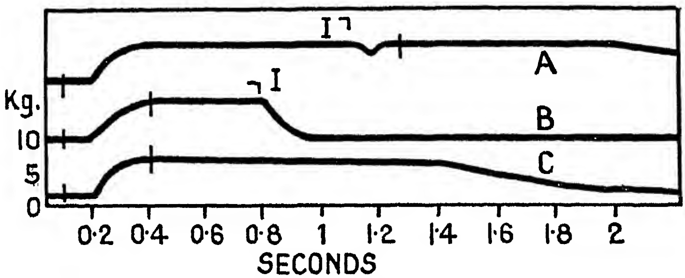


Fig. 338.—Inhibitory Distinction between Stimulation Plateau and After-discharge Plateau of Crossed Extensor Reflex. (Liddell and Sherrington, *Proc. roy. Soc. B.*, 1925.)

Vastocrureus. Between the vertical lines on the records, stimulation of afferent nerve to elicit crossed extensor reflex; I, delivery of single break shock to inhibitory afferent.
A, delivery of inhibition in stimulation plateau; B, delivery of inhibition in after-discharge plateau; C, repetition of reflex B with no inhibitory shock delivered (isometric records). Note duration of after-discharge.

stimulated together. In brief, occlusion is due to, and is convincing evidence of, afferent fibres overlapping in their central distribution (Fig. 339, A).

(12) Subliminal Fringe.—Sometimes, however, the tension yielded by *a* and *b* combined is *greater* than the sum of the two reflex responses taken singly (Fig. 339, B). This result indicates that each afferent while fully activating a certain number of motor neurones acts also on a further number *subliminally*, and that some of these subliminally influenced motor neurones are common to *a* and *b*. Concurrent subliminal excitations can thus sum to produce liminal stimulation.[1] This type of result is another illustration of *spatial* summation.

Subliminal fringe is of great importance in reflex co-ordination. It enables one level in the nervous system to reinforce the action of another. Thus, sometimes, feeble stretch of an extensor muscle (p. 588) produces a weak response; rotation of the head alone (p. 591) may also produce little increase in the activity of the muscle. But the combination of the two procedures may give rise to a considerable contraction of the muscle—again an example of summation of subliminal fringes.

[1] Denny-Brown and Sherrington, *J. Physiol.*, 1928, *66*, 175.

(13) Motor Neurone Discharge Rate and Rate of Afferent Stimulation.—There need not necessarily be any regular relationship between the rate of stimulation of an afferent nerve and the discharge rate of the motor neurone. The latter will depend on the degree of summation and after-discharge that occurs. A single afferent volley may produce a *repetitive* discharge from the motor neurones (because of after-discharge); similarly, each volley in a train of volleys may generate several motor impulses. The discharge rate would then be considerably higher than the afferent stimulation rate. On the other hand, discharge may not occur till a number of afferent volleys has reached the centre and effective summation has occurred; the discharge rate would then be a fraction of the afferent stimulation rate.

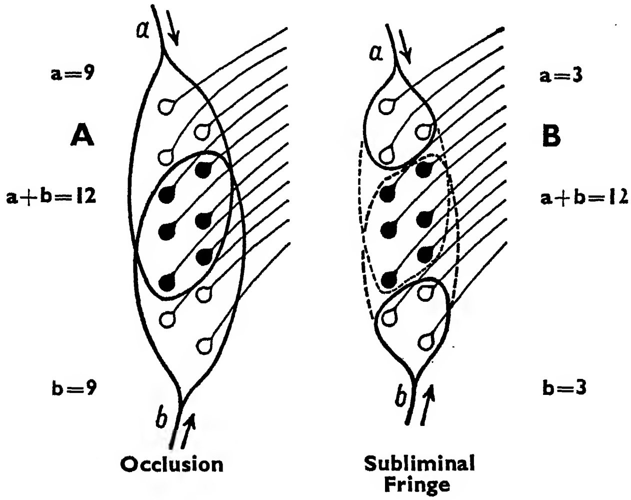


Fig. 339.—Occlusion and Subliminal Fringe.

The diagram shows a group of motor neurones. The clear cells are exclusively influenced by the afferent fibres *a* or *b*; the dark cells are common to both *a* and *b*.
A: Stimulation of *a* or *b* effectively excites 9 motor neurones. Stimulation of *a* and *b* together excites only 12 motor neurones, because 6 are common to both afferents. Occlusion takes place.
B: Stimulation of *a* or *b* effectively excites 3 motor neurones (enclosed by continuous line) and produces a subliminal effect on another 6 motor neurones (enclosed by dotted line). Stimulation of *a* and *b* stimulates 12 motor neurones as the two subliminal fringes effectively sum.

Thus in one experiment, stimulation at 3, 88, and 228 per second elicited discharge rates of 16, 14, and 18 per second. During a reflex like the flexor reflex, the discharge rate of the motor neurones may not be uniform; it may initially rise rapidly to a maximum and then speedily decline.

Increasing the *frequency* of afferent stimulation increases the tension of the reflex response; as more afferent impulses arrive in unit time there are greater opportunities for central summation, more motor neurones are effectively stimulated and they may discharge at a higher frequency.

Many of the features of reflex excitation in the somatic nervous system (*e.g.* summation, irradiation, recruitment, after-discharge) can also be observed in reflexes in the autonomic nervous system.

(14) Gradation of Central Excitation.—(i) It may be subliminal, *i.e.*

there is no motor neurone discharge, but there is opportunity for summation to occur (p. 533).

(ii) The motor neurone may discharge once and a simple twitch results.

(iii) The motor neurone may discharge repetitively at a low frequency; the corresponding group of muscle fibres goes into partial tetanus (subtetanus).

(iv) When the motor neurone discharges at about 60–90 times per second it is said to be excited *maximally*, as with such a discharge rate full tetanus of the corresponding motor unit usually results (p. 500).

(v) The motor neurone may, however, discharge at still higher rates, up to 100 or 150 times per second; this is called *supramaximal* activity. As far as the muscle is concerned this high rate of discharge is wasteful, because no greater mechanical response develops; the high discharge rate can, however, be seen in the muscle action potentials. A centre driven at a high rate is protected to some extent from inhibitory influences.

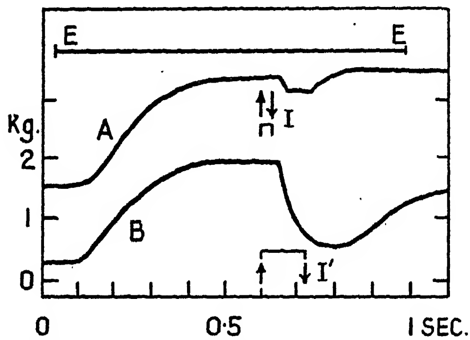


FIG. 340.—Recruitment and After-Discharge in Inhibitory Reflexes. (Liddell and Sherrington, *Proc. roy. Soc. B.*, 1925.)

Knee Extensor. In both records from E to E, stimulation of contralateral afferent to produce crossed extensor reflex. During the period indicated by the arrows (I, I′), stimulation of ipsilateral inhibitory afferent, for a shorter time in A, and for a longer time in B. Note that *prolongation of inhibitory stimulation increases the extent of the inhibition*; following it is a horizontal plateau which is interpreted as representing *inhibitory after-discharge*; *central inhibition persists and prevents for 0·05 sec. the redevelopment of tension in the muscle.*

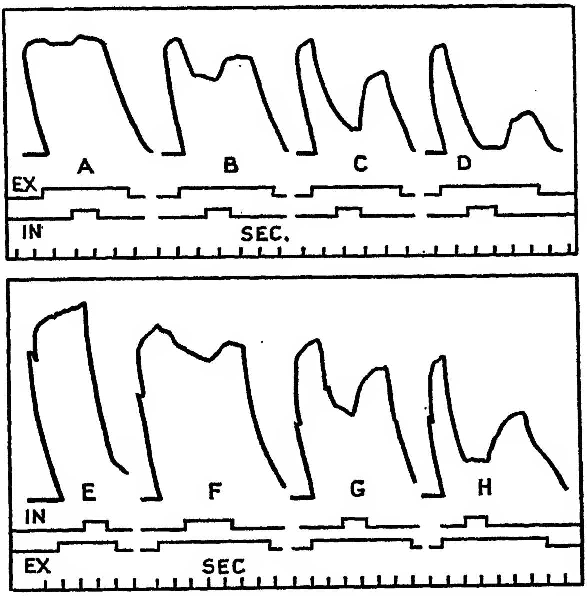
By means of its *wide range of discharge rate each motor neurone can finely grade the activity of the muscle fibres which it supplies.*

(vi) Finally, an afferent nerve, by stimulating *varying numbers* of motor neurones in the "pool" under its possible control, can regulate the *number* of motor units in action, and consequently the tension which is developed by the muscle.

General Features of Reflex Inhibition.—Reflex inhibition resembles reflex excitation in many respects; the fundamental difference is that the impulses arriving at the synaptic terminals on the motor neurones *depress* (instead of stimulating) the cells. An *inhibitory afferent* is one which, when appropriately stimulated, depresses the group of motor neurones under examination, *i.e.* it inhibits the muscle fibres supplied by these motor neurones. Inhibition can, generally, only be demonstrated against a *background of activity*; the background classically used in the study of reflex inhibition is the discharge of the spinal extensor neurones set up by stretch reflexes in decerebrate rigidity (p. 584) or by the crossed extensor reflex.

The characteristics of reflex inhibition are:

(1) DELAY: partly peripheral, partly central (cf. p. 532). The central delay depends on the number of internuncials traversed (transmission through these is of course excitatory). The synaptic delay at the motor neurone at which inhibition is produced, is about 1 m.sec.: during this time the impulse at the synaptic terminals liberates the hypothetical inhibitory chemical transmitter, which sets up the supposed, positive localized electro-



FIGS. 341 and 342.—Antagonism of Central Excitation and Central Inhibition. (After Sherrington, *Proc. roy. Soc. B.*, 1909, *81*, 249.)

Records of reflex contraction of a flexor muscle.
EX: at each rise of the signal, stimulate the ipsilateral excitatory afferent nerve which elicits reflex contraction of the flexor muscle.
IN: at each rise of the signal, stimulate the contralateral afferent nerve which inhibits the flexor motor neurones.
A, B, C, D: strength of excitatory stimulus kept constant; strength of inhibitory stimulus progressively increased.
E, F, G, H: strength of inhibitory stimulus kept constant: strength of excitatory stimulus progressively decreased.

tonus which, on reaching threshold value, depresses the "firing" of the motor neurones.

(2) SUMMATION.—As with excitation, central inhibition may be subliminal; in terms of the working hypothesis, the positive electrotonus does not reach threshold value. Both spatial summation (produced by simultaneous stimulation of two inhibitory afferents) and so-called temporal summation (produced by repetitive stimulation of a single inhibitory afferent) can be produced (cf. discussion on p. 533).

(3) FATIGUE.—If an inhibitory reflex is elicited at short intervals, the inhibitory response (*i.e.* the degree of muscular relaxation) becomes smaller;

afferent (IN) acting on the same group of motor neurones. By suitably adjusting the strength of stimulation of afferents EX and IN, the activity of the motor neurones can be finely graded (Fig. 341). With increased strength of afferent IN, the number of motor neurones which are discharging is progressively reduced. In the case of an individual motor neurone, its discharge rate can be slowed down from supramaximal (150 per sec.) to maximal (60–90 per sec.), then to lower, subtetanizing, frequencies; finally the discharge is completely stopped (Fig. 343).

Reciprocal Inhibition (Reciprocal Innervation).—Reflex inhibition of the antagonist muscles is an almost invariable accompaniment of reflex excitation of the protagonist muscles. This is the principle of *reciprocal innervation of muscles*. Movements are thus facilitated, as relaxation of antagonists proceeds *pari passu* with contraction of protagonists. This important general law can be readily demonstrated (Fig. 344).

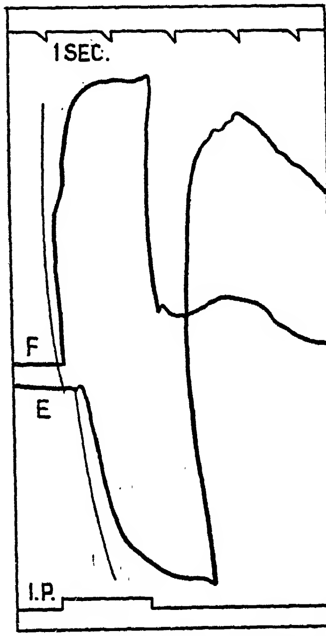


FIG. 344.—Reciprocal Inhibition of Antagonistic Muscles. (Sherrington, *Quart. J. exp. Physiol.*, 1913.)

F, semitendinosus, a knee flexor; E, vastocrureus, a knee extensor. During rise of signal line IP, stimulate ipsilateral popliteal nerve. The inhibition of the extensor is accompanied by a simultaneous contraction of the flexor muscle. Myograph writer for E is set a little to the right of that for F, so that ascent of F and descent of E are in fact practically synchronous. Relaxation of the extensor is followed by a marked *rebound* contraction above the original level.

In a decerebrate animal the flexors of the knee and their antagonist the vastocrureus may be detached from their insertions and attached to levers. If a nocuous stimulus is applied to the sole of the foot to elicit the flexor reflex, it is found that as the flexors contract the extensors elongate. It can be shown that the two processes are simultaneously induced. The receptive field, the nature of the stimulus, the latent period and the number of subliminal stimuli necessary to produce summation are identical for contraction of the hamstrings and for relaxation of the vastocrureus. The afferent impulses which enter the spinal cord stimulate the motor neurones supplying the protagonists and inhibit the motor neurones supplying the antagonists (Figs. 344 and 345).[1]

Fig. 344 is analysed in detail in the Appendix, p. 1123, where the subject of reciprocal innervation is discussed very fully.

Reciprocal innervation can also be demonstrated in *postures* (p. 590), following stimulation of the *motor cortex* (p. 634) or of the *frontal eyefield* (p. 638), following *cerebellar* stimulation, in labyrinthine *nystagmus* (p. 601)

[1] Sometimes, however, it is more advantageous to have simultaneous contraction of protagonists and antagonists—as when it is desired to maintain a rigid posture (p. 590).

and in *stretch reflexes, e.g.* the *knee-jerk* (p. 644). To some extent the principle also applies to the autonomic nervous system.

Reciprocal innervation in *voluntary movement* is fully discussed on pp. 649 *et seq.*

Central and Peripheral Inhibition.—In *central* inhibition the seat of the inhibitory process is in the centre, *e.g.* inhibition of extensor motor neurones

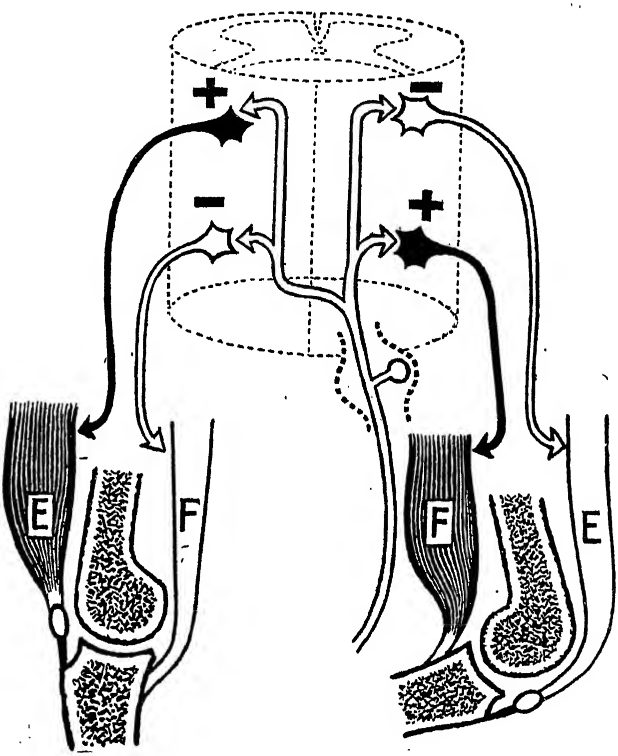


FIG. 345.—Diagram illustrating Mechanism of Reciprocal Innervation. (Modified from Sherrington, *Integrative Action of Nervous System.*)

E = Knee extensors. F = Knee flexors. + = Central excitation. — = Central inhibition.
Stimulation of central end of afferent nerve on right side reflexly produces flexor reflex on right side and crossed extensor reflex on left side. In the case of each reflex response, contraction of protagonists is accompanied by reciprocal inhibition of antagonists.

n the flexor reflex; inhibition of the vasomotor centre by the aortic or sinus nerve (p. 739); inhibition of the inspiratory centre by vagal afferents from the lungs (p. 388). In *peripheral* inhibition the seat of the inhibitory process is in the peripheral tissues, *e.g.* vagal inhibition of the heart (p. 268), sympathetic inhibition of the small intestine. Suppression (p. 620) and conditioned inhibition (p. 679) are special types of central inhibition.

THE AFFERENT NERVOUS SYSTEM[1]

Skin and Muscle Receptors.—These may be classified as follows:

(i) In *Muscle and Tendon*: (*a*) Muscle spindle; (*b*) Organ of Golgi.

(ii) *Organized Skin Receptors:* (*a*) Pacinian corpuscle; (*b*) Tactile corpuscle; (*c*) End bulbs.

(iii) *Pain Receptors:* (*a*) free nerve endings; (*b*) plexiform networks.

(1) Musculo-Tendinous Endings.—(i) The *muscle spindle* is a fusiform body lying between the muscle bundles, with its long axis parallel to them. It consists of 6–12 poorly differentiated thin "intrafusal" muscle fibres rich in nuclei and sarcoplasm; the transverse striations are not well marked. The spindle has a complex innervation: (*a*) An afferent nerve fibre enters the spindle and breaks up into non-medullated branches, each of which winds in a spiral or annular manner round a single intrafusal fibre ("annulo-spiral ending") (Fig. 346). These receptors respond (α) to passive stretch, and (β) to changes in the tension of the intrafusal fibres resulting from their active contraction. When the muscle fibres *outside* the spindle contract, the degree of stretch of the spindle is generally decreased but sometimes is increased. (*b*) The intrafusal fibres receive a *motor* innervation from thin medullated somatic fibres.[2] (*c*) A non-medullated afferent fibre; this may be an associated pain receptor.

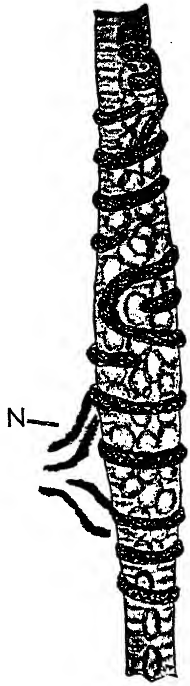


Fig. 346.—Intrafusal Fibre of Muscle Spindle and its Nerve Supply (J. Boeke from Sharpey-Schafer, *Essentials of Histology*).
N = Afferent nerve supplying the muscle fibre. Note the annular method of ending of the nerve fibre round the highly nucleated intrafusal fibre.

(ii) The *organ of Golgi* is found in tendon close to its point of attachment to the muscle. The tendon fibres separate into a number of small bundles; nerve fibres penetrate between the fasciculi, and their medullary sheaths stop short; the axons end in terminal arborizations beset with irregular varicosities (Fig. 347). These organs are also found in the connective tissue between the muscle fibres. The Golgi organ is a stretch receptor; it is always stretched when the muscle contracts.

(2) Organized Skin Receptors.—These receptors which subserve the sensations of touch, light pressure, heat, and cold are constructed on a uniform plan. They consist of an outer lamellated connective-tissue capsule, and a core of soft nucleated cells within which the axon, having lost its medullary sheath, ends simply or as an arborescence. The receptors to be described vary chiefly in the complexity of their design.

(i) *Pacinian corpuscles* are large receptors found in the subcutaneous tissues of the hands and feet and in the neighbourhood of tendons and joints. The capsule consists of a number of concentric fibrous coats arranged like the layers of an onion; the soft core is cylindrical in shape. The

[1] Head, *Studies in Neurology*, 2 vols., 1920. Stopford, *Sensation and Sensory Pathway*, London, 1930; Symposia, *Res. Publ. nerv. ment. Dis.*, Sensation, 1935; Pain, 1943. Newton, Cutaneous Sensation, *Recent Advances in Physiology*, 7th edn., London, 1949.

[2] Hunt and Kuffler, *J. Physiol.*, 1951, *113*, 283, 298.

nerve fibre passes down the middle of the core to its farther end to form a terminal arborization. The Pacinian corpuscles respond to deformation.

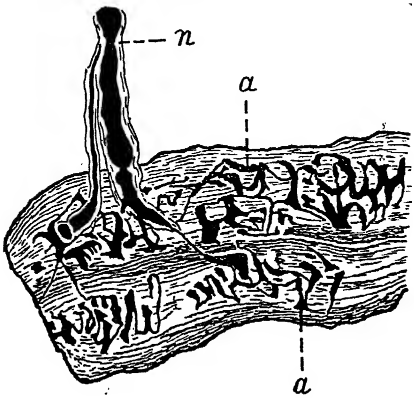


Fig. 347.—Organ of Golgi. (Ciaccio, from Sharpey-Schafer, *Essentials of Histology.*)

n = Afferent nerve fibre. a = Arborization of axis cylinder between the tendon-bundles.

(ii) *Tactile corpuscles* (of Meissner) are found in the papillæ of the skin. These touch receptors are ellipsoidal in shape. The afferent nerve fibre enters, ends in a complex ramification in the soft central core of the corpuscle (Fig. 348). The Meissner corpuscles always occur in groups of two or three (never singly); in finger skin, about ten such groups are found in an area of 1 sq. mm. Other tactile receptors are found forming a basket-like ending round the base of the *hair-follicles* (Fig. 349); their arrangement suggests that they would be stimulated by mechanical displacement of the hair.

(iii) *End bulbs* (of Krause) occur in the conjunctiva, the papillæ of the lips and tongue, the sheaths of nerve trunks, the skin of the genital organs in both sexes, and in the neighbourhood of joints. They are *spherical* in shape (Fig. 349), but otherwise resemble the tactile corpuscles. They are believed to respond to *cold.*

A somewhat simpler type of ending (the organ of Golgi-Mazzoni (Fig. 349)) is said to be specifically responsive to *warmth.*

(3) Pain Receptors.—The receptors are free (unencapsulated) thin nerve filaments both non-medullated and medullated. The muscle spindles and many of the organised skin receptors receive an "accessory" non-medullated fibre which is believed to be a pain receptor (Fig. 348). In the skin, two nerve plexuses are present which are related to pain sensibility: (i) a

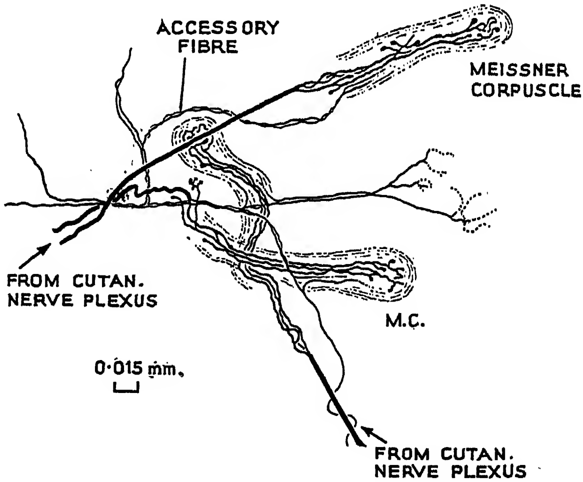


Fig. 348.—Meissner's (Touch) Corpuscles. (Weddell, *J. Anat.*, 1941, *75*, 441.)

Three end organs, each supplied by a separate medullated nerve fibre were found in an area 0·15 mm. wide. Note carefully the accessory *non-medullated* nerve fibre supplying the corpuscles; it may mediate *pain* sensibility.

superficial plexus under the epidermis, from which filaments pass to end between the epithelial cells or actually in the cytoplasm of the cells. These free endings can be readily acted on by chemical changes in the cells produced by noxious agents; (ii) a *deeper* subepidermal plexus. Free nerve endings are present round the blood vessels and in the aponeurotic sheaths of skeletal muscle and probably subserve pain in these localities.

Terminology.—Afferent impulses may be divided into—

(i) *Exteroceptive:* those set up by events in the outer world, such as auditory, visual, or olfactory stimuli, or those giving rise to sensations of touch, pain, or temperature.

(ii) *Proprioceptive:* coming from the muscles and adjacent deep structures (*e.g.* ligaments, tendons, joints) (also called "kinæsthetic impulses") and from the labyrinths. Proprioceptive impulses give information about the position of the head and the various parts of the body at rest and during movement.

(iii) *Enteroceptive:* arising from the viscera.

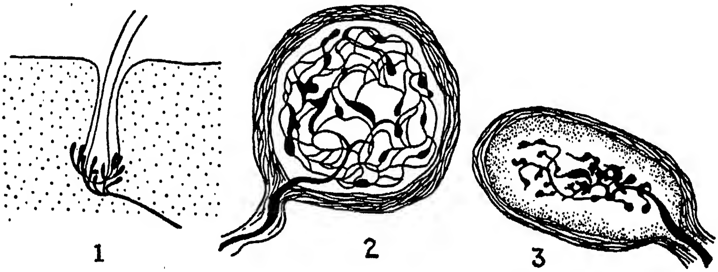


FIG. 349.—Various Forms of Cutaneous Sensory Nerve Endings. (Fulton, *Physiology of Nervous System*, N.Y.)

(1) Basket-like ending round the base of a hair follicle.
(2) End-bulb (? subserves cold sensation).
(3) Organ of Golgi-Mazzoni (? subserves warmth sensation).

General Properties of Receptors.[1]—(1) METHODS OF STUDY.—Using amplifiers and the cathode ray oscillograph it is possible to record the action potential (which accompanies the nerve impulse) in *single* nerve fibres and thus get information of great interest. It is of less value to study the action potentials in a nerve trunk which contains many fibres: the action potentials are *out of step* in adjacent fibres, and the deflections recorded are simply the algebraic sum of the electrical changes (at any moment) in *all* the fibres and do not represent what happens in any *individual* fibre. It is usually necessary (i) to reduce the number of conducting fibres in the nerve under examination to very few (if possible to a single one) by appropriate section, or (ii) to simplify the peripheral end by having only a single sense organ in action. The latter procedure is well demonstrated in the case of the minute sterno-cutaneous muscle of the frog which contains several *muscle spindles*. When the whole muscle is stretched all the endings discharge asynchronously and numerous irregular impulses are recorded in the nerve (Fig. 350, A). Successive parts of the muscle are cut off, till one spindle only is left: stretch of the

[1]Adrian, *Basis of Sensation*, London, 1928; *Physiol. Rev.*, 1930, *10*, 336.

muscle now gives rise to a slow regular series of impulses from the single ending (Fig. 350, B). The *Pacinian corpuscle* on the other hand is sufficiently large to be isolated and stimulated directly.

(2) Specificity of Response. Peripheral Analysis.—The receptors (except those for pain), because of their specialized structure respond only when the appropriate, specific stimulus is applied ; other kinds of stimuli are ineffective. Thus the tactile corpuscles or Pacinian corpuscles respond to deformation ; the tendon organs to stretch ; the otolith organ to the pull of gravity ; the retina to light ; the cochlea to vibrations of the basement membrane set up by sound. The receptors are thus *peripheral analysers* which respond to a specific environmental change by generating nerve impulses.

(3) Repetitive Response. Adaptation.—As the skin and muscle

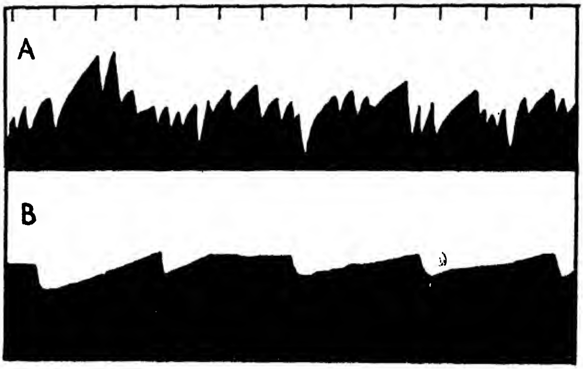


Fig. 350.—Action Currents set up by Stretch Receptors in Muscle. (Adrian and Zotterman, *J. Physiol.*)

Record of afferent impulses in nerve from stretched muscle. Time, 0·01 sec. Capillary electrometer (amplification × 490).
A, Intact muscle, 2 g. weight applied. Numerous irregular impulses.
B, Muscle divided till it contained a single nerve ending ; 1 g. weight applied. Slow regular series of impulses.

receptors are in essence nothing but appropriately arranged free or encapsulated, non-myelinated nerve fibres, we must consider first the response of a nerve fibre to stimulation. If a constant stimulus in the form of a constant current is applied to a nerve fibre, a single impulse is set up at the moment the current is made (Fig. 351). The fibre immediately becomes refractory and then recovers to normal, but it does *not* respond again to the *presence* of the current which is passing through it, though the first application of this current proved an adequate stimulus. This phenomenon is described as " accommodation " (p. 494) : *i.e.* the constant current ceases to be an effective stimulus.

Now let us repeat this procedure with receptors, *i.e.* the appropriate natural stimulus is applied and maintained. The receptor responds to a constant stimulus not with a single discharge but with a repetitive burst of nerve impulses ; accommodation or *adaptation* as it is called in the case of the receptors, is not immediate. The duration of the burst varies greatly among the different receptors (Fig. 351)

(i) In the case of stretch receptors, *e.g.* the muscle spindles, adaptation takes place very slowly; the discharge continues for as long as the muscle is stretched, though the *frequency* of the discharge tends to decline (cf. p. 589). Stretch receptors with practically identical characteristics are found: (*a*) in the otolith organs (p. 598), (*b*) in the vaso-sensory zones (aortic arch and carotid sinus, Fig. 479, p. 742), (*c*) in the alveoli of the lungs (Fig. 231, p. 387). The behaviour of the receptors in the ampullæ of the semicircular canals is discussed on p. 599. Stretch receptors are also present in the hollow viscera, *e.g.* stomach, intestine, and urinary bladder (p. 770).

(ii) In *touch* receptors adaptation is rapid. Thus if a hair is bent, impulses are set up only during the movement, but not after it has ceased, although the hair is kept in the abnormal position. Similarly, if hairless skin is touched the discharge may only last for 0·2 second.

(iii) The rate of adaptation in *temperature* receptors is intermediate between that of muscular and tactile receptors.

(iv) *Pain* receptors are considered on p. 553. Some pain receptors, *e.g.* those in the cornea, show little adaptation.

Under natural conditions adaptation may be less of an impediment than appears from controlled experiments. The natural stimulus is usually of a more *variable* character, not affecting exactly the same group of receptors all the time; some of the receptors in the stimulated area can thus recover during brief quiescent periods (free from stimulation) and respond maximally again the next time the stimulus affects them.

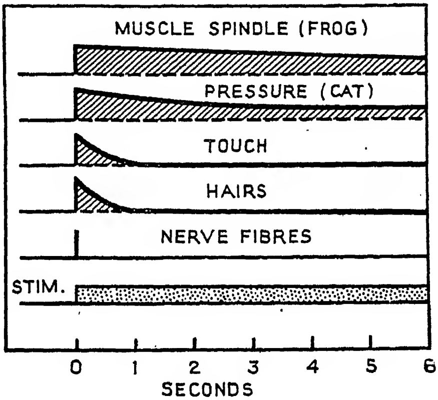


FIG. 351.—Rate of Adaptation of Nerve Endings. (After Adrian, *Basis of Sensation*, 1928.)

Height of curves is proportional to frequency of discharge of nerve ending at various times after applying the stimulus.

(4) EFFECT OF STRENGTH (INTENSITY) OF STIMULUS.—As the strength of stimulation is increased the frequency of the discharge rate rises from low levels, *e.g.* 5–10 per second to maximal levels of 300 per second or higher. Thus in Fig. 352, pressures of 250 g. and 500 g. applied to the cat's toe-pad, produced peak discharge rates of about 250 and 400 per second respectively. A stimulus which increases slowly in intensity produces a lower peak discharge rate than a stimulus that rises very rapidly to the same maximum (Fig. 353). In all receptors,i ntensity of stimulation is translated into impulse frequency, *i.e.* the numbers of impulses per second transmitted along the related afferent fibre.

(5) EFFECT OF EXTENT OF STIMULATION.—An extensive stimulus activates more receptors and consequently impulses pass back along a larger number of nerve fibres. (Similarly, increasing the strength of stimulation which is applied to a nerve trunk brings into action more nerve fibres.)

(6) ACTION POTENTIAL.—The action potential (and presumably the

associated nervous impulse) varies with the diameter and type of conducting afferent fibre, but not with the nature of the receptor. The quality (*modality*) of a sensation (*e.g.* whether it is touch, heat, cold) thus in no way depends on the characteristics of the impulse in the afferent nerves.

(7) The factors concerned in determining the *localization* and *quality* of the sensation are discussed on p. 570.

Temperature.—A substance which has the same temperature as the skin produces no sensation of temperature. When a suitably warmed instrument is passed over the skin vivid appreciation of warmth occurs at certain points; these are the "warm" or "hot" spots; similarly, "cold" spots may be mapped out. The two types of spot do not coincide with

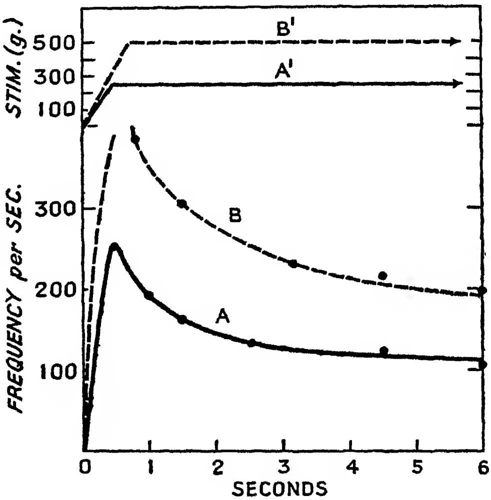


FIG. 352.—Relation of Strength of Stimulus and resulting Frequency of Afferent Impulses. (Modified from Adrian, *Basis of Sensation.*)

Stimulus A′: weight of 250 g. applied to cat's toe-pad. Resulting maximal frequency (A) is 250 per sec.
Stimulus B′: weight of 500 g. Resulting maximal frequency (B) exceeds 400 per sec.
Note rapid adaptation in both cases as stimulus is maintained.

each other and they do not respond to mechanical or electrical stimulation. The "cold" spots are far more numerous than the "warm" spots. They are mostly found on the chest, nipples, nose, anterior surface of the arms and forearm and on the abdomen. They are less numerous in exposed parts such as the face, hands, and mucous membranes.

The temperature receptors show a special type of adaptation which is demonstrated by the three basins experiment. One hand is plunged into hot water and the other into cold water for a few moments: both are then placed in a basin of tepid water. The former gives rise to a sensation of cold, the latter to one of warmth.

Touch.—This sensation is studied by means of Frey's hairs, which are made of varying cross-section; when bent they exert a pressure which

varies directly with their thickness. The touch spots bear no relation to the temperature spots. The adequate stimulus for touch spots is deformation of the skin by pulling or pushing; uniform pressure (*e.g.* dipping the finger into a beaker of mercury) is ineffective. Touch spots are most numerous on the finger-tips, and very sparse on the shoulders and the shins. The most sensitive areas of the body are the tongue, nose, lips, and finger-tips. The touch spots adapt rapidly; thus the continual stimulus of clothing is not appreciated, because adaptation soon takes place (see p. 550). The *hair follicles* have a rich sensory innervation. A slight touch to the hair acts in the same way as a stimulus applied to the long arm of a lever—the short arm of the lever being the root of the hair, which is inside the skin. The stimulus is thus multiplied about fivefold during transmission through the

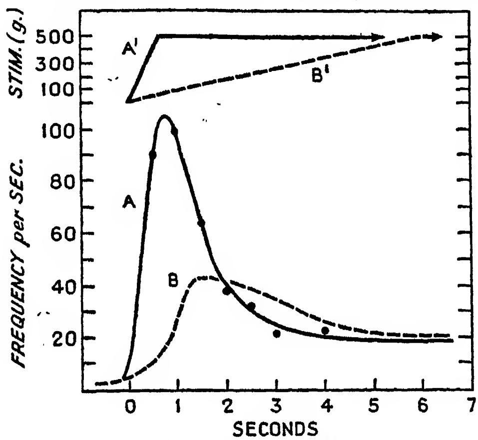


FIG. 353.—Relationship between Rate of Stretch of Muscle and Frequency of Impulses set up in Receptors. (Modified from Adrian, *Basis of Sensation.*)

Vertical Axis—*Above:* Weight attached to muscle to produce stretch and stimulate receptors. *Below:* Frequency of afferent nerve impulses per sec. set up by stretching muscle and recorded in distal end of cut muscle nerve.

In A′, the stretch reached maximum in less than 1 sec.; in B′ the same maximum stretch was attained after 6 sec. Note that the maximal frequency of the nerve impulses (A) in response to A′ is over 100 per sec. compared with 40 per sec. (B) in response to B′.

hair. Consequently, shaving the skin considerably diminishes its sensibility. The sensation of *light superficial pressure* probably also depends on the response to touch receptors.

Pain Sensibility.—(1) SPECIFIC QUALITY OF PAIN.—There is evidence to show that pain is a specific quality (modality) of sensation and does not merely represent the result of excessive stimulation of any kind of receptor.

(i) When some tissues, *e.g.* the cornea, are stimulated they only give rise to pain; the only nerve endings in the cornea are naked fibres that pass between the epithelial cells.

(ii) When an interrupted jet of air is applied to the skin it may cause an impulse discharge at a frequency as high as 300 per second, but the animal, judged by its behaviour, experiences no pain.

(2) STIMULUS FOR PAIN.—Unlike the specialized receptors, the pain endings can be stimulated by a wide range of stimuli which only have in

common the property of *damaging* tissues, *i.e.* they are nocuous, harmful stimuli. Thus excessive heat or cold, excessive stretch or tension, many chemical agents (products of ischæmia, acid, alkali, hypertonic solutions), electric currents, scratching, or cutting may produce pain. It is possible that the effective stimulus to the pain endings is some chemical agent released by injury.[1]

(3) FIBRE TYPES INVOLVED.—Nocuous stimulation sets up a prolonged discharge of impulses, with little sign of adaptation, conducted in two kinds of fibres: (*a*) very fine non-medullated C fibres; the action potentials in these fibres are, characteristically, of low voltage and have slow time relations; (*b*) thin medullated A fibres of larger diameter than the C fibres. The notion of a *double* peripheral pain path is supported by the following evidence:

(i) If a hot body at a temperature of 60–65° C. is applied to the skin for 0·3 sec., a "double" pain is felt; there is a "former" and a "latter" pain, generally called the "fast" and "slow" pain, which are separated

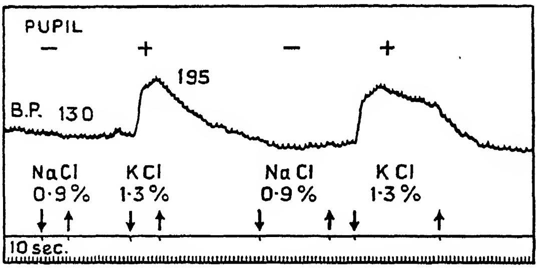


FIG. 354.—Stimulation of Pain Afferents in Dorsal Nerve Roots produces Reflex Rise of Blood Pressure and Dilatation of Pupil. (Calma and Wright, *J. Physiol.*, 1947, *106*, 211.)

Records from above downwards:
Pupil diameter: —constricted; +dilated; arterial blood pressure in mm. Hg; signal line; time in 10 sec.
Apply solutions directly to exposed lumbar dorsal nerve root.
Isotonic (0·9%) NaCl, no effect. Isotonic KCl (1·3%) stimulates pain afferents: as a result there is a reflex rise of blood pressure and dilatation of pupil.

by a distinct interval; this observation suggests the presence of two kinds of pain fibres with different rates of conduction. In conformity with this view the interval between the "fast" and the "slow" pain is longer if the stimulus is applied to the foot and shorter if it is applied to the hand.

(ii) In the injury to the dorsal nerve roots which occurs in tabes (p. 561), "fast" pain may be lost and only the "slow" component persists, presumably because the disease abolishes conduction in the thicker fibres first. The time interval between the application of a stimulus to the knee and the resulting sensation of "slow" pain is about 1 sec.; it is about 2 secs. when the foot is stimulated. The difference in time corresponds to a conduction rate in the nerves involved of under 0·5 metre per sec. (*i.e.* that of C fibres).[2]

(iii) Local application of procaine paralyses the finest nerve fibres first and the thickest last; asphyxia (induced by applying a cuff inflated to above systolic pressure to the arm) blocks conduction (under the cuff) in the thickest

[1] The pain sense attaches to a tissue "a specific sense of its own injuries."

[2] It should be pointed out, however, that the latency in some cases may be as long as 20 secs.!

fibres first and in the thinnest last. As might be expected, procaine abolishes "slow" pain first, while asphyxia abolishes "fast" pain first.

(iv) The non-medullated and fine medullated fibres in the dorsal nerve roots constitute the lateral division of the root; they enter the substantia gelatinosa (at the apex of the dorsal grey horn), where they end. Stimulation of the dorsal root produces reflexly an increase in the rate of the heart, a rise of blood pressure (Fig. 354) and increased respiration, *i.e.* the changes which are generally associated with nocuous stimulation; these reflex reactions disappear if the lateral division of the root is severed.

(4) Complex Pain.—The specific stimuli for certain receptors, when intense, may stimulate adjacent pain fibres; thus excessive heat, cold, or pressure arouses pain as well as the specific sensation. Similarly if the skin is stroked vigorously with a wooden pin, there is initially a response from the receptors for touch and light pressure; when a sensation of burning pain develops, impulses can be demonstrated travelling in C pain fibres. A complex sensation can thus be set up by natural stimuli as different groups of receptors are stimulated simultaneously or successively; the exact character of the sensation will depend on the variety of receptors stimulated.

(5) Affective Component of Pain.—Pain not only represents a special quality (modality) of sensation, but it is characterized by a large emotional (*affective*) accompaniment: pain is always unpleasant, and may become intolerable. By contrast, some kinds of sensation are almost devoid of affect, *e.g.* muscle sense; moderate temperature may be associated with a pleasurable affect; touch may be associated with every degree of affect, according to what is being touched (and by whom).

(6) Localization of pain and *referred pain* are considered on pp. 571, 756 (Ischæmic pain is discussed on p. 750, visceral pain on pp. 752 *et seq.*

(7) Prepotence of Reflex Effects produced by Nociceptive Impulses.—Impulses set up by nocuous stimuli, in addition to arousing the sensation of pain, also set up widespread *reflex* effects, even in spinal animals in which, of course, pain sensation is non-existent. It is convenient to call nerve impulses set up by nocuous stimuli, *nociceptive* impulses; "what from the point of view of sensation are pain nerves, are from the point of view of reflex reactions conveniently termed nociceptive nerves." Nociceptive impulses produce widespread reflex reactions; *e.g.* a needle prick to the sole of the foot in the spinal animal produces the flexion reflex; stimulation of nociceptive nerves in the dorsal nerve roots sets up marked reflex changes in circulation and respiration (*supra*). Nociceptive impulses from viscera produce reflex contraction of muscles and alterations in glandular secretion and in vasomotor tone. When there is competition between reflexes, those which are set up by nociceptive impulses "as a rule dominate with peculiar certainty and facility," *i.e.* they are *prepotent* reflexes. Thus when reflex competition is taking place, the nociceptive flexor reflex overcomes the proprioceptive extensor thrust reflex or the proprioceptive reflex which is responsible for quadriceps tone in decerebrate rigidity.[1]

[1] *Sexual* reflexes (not less important to the species than the nociceptive, and like the latter also associated in the intact animal with an intense affect) also show a high degree of prepotence. Thus the *clasp reflex* of the male frog, which is set up by the female coming in contact with the skin in the sternal and adjacent region, cannot be inhibited by severe mutilation of the limbs or of internal organs, or by strong stimulation of the sciatic nerve; it is, however, sometimes depressed by strong nocuous stimulation of the sternal skin.

Conduction of Afferent Impulses in the Peripheral Nerves.[1]—The afferent impulses in the peripheral nerves are conducted partly in (so-called) motor nerves and partly in cutaneous sensory nerves.

(i) The former convey impulses from the deep structures, such as the muscles, tendons, ligaments, periosteum, and joints. The so-called " motor " nerve to a muscle is really a mixed efferent and afferent nerve ; 40% of the fibres in a " motor " nerve are afferent and have their cell bodies in the dorsal root ganglia. These deep afferents connect with cells in the spinal cord, brain stem, cerebellum, and cerebrum and reflexly modify posture and movements. Some of these fibres after relaying, ultimately reach the post-central cortex where the impulses give rise to several kinds of sensation, *e.g.* deep pressure, sense of position, sense of passive and active movement and vibration sense in bone (these sensations are referred to collectively as *deep sensibility*).

(ii) The cutaneous nerves convey impulses which, (*a*) produce reflex effects and (*b*) on reaching the cerebral cortex give rise to sensations of touch, heat, cold, and pain (*cutaneous sensibility*).

Effects of Cutaneous Nerve Section.[2]—IMMEDIATE EFFECTS.—In Head's classical experiments, the radial and external cutaneous nerves were exposed in the neighbourhood of the elbow and small portions were excised ; the cut ends were at once united to facilitate regeneration of the fibres. All forms of cutaneous, *i.e. superficial sensibility* were abolished over the radial half of the forearm and the back of the hand. Deep sensibility, however, was retained as it is mediated by afferent fibres in the motor nerves. Deep pressure was recognized. Excessive pressure gave rise to aching pain. There was awareness of the position and of movements of the part.

Recovery of Skin Sensibility.—(1) ANATOMICAL CONSIDERATIONS.—The histological changes of degeneration and regeneration following nerve section are described on p. 496 ; broader anatomical factors must now be discussed. Under ideal conditions each sprouting central axon would grow down its old neurilemmal sheath to its own type of receptor in its original locality. Such perfect anatomical recovery never takes place even when the two cut ends of the nerve are carefully stitched together immediately after the section, because it is certain that the two cut ends of *any one fibre* will never[3] be placed in apposition ; much better anatomical repair may occur, for obvious reasons, when a nerve is *crushed* and not cut. When the two ends of a cut nerve are separated by a gap, many of the sprouting central axons get lost, never reach the peripheral sheaths, and so never function. Such fibres as do reach the distal end of the nerve will make one or more errors in the connections they establish.

(i) They may connect up with a receptor which is of their own type but which lies in a different region of the skin, *e.g.* a touch fibre originally supplying the base of the thumb may connect with a receptor at the tip of the thumb.

(ii) They may connect with a different type of receptor, *e.g.* a touch fibre may connect with a temperature receptor.

(iii) Mistakes (i) and (ii) will often be combined.

[1] Walshe, *Brain*, 1942, *65*, 48–112.

[2] If a mixed nerve which supplies muscles as well as skin is cut, there is in addition paralysis of the corresponding muscles of the lower motor neurone type (cf. p. 505) and loss of deep sensibility.

[3] " What, never ? Well, hardly ever."

(iv) A fibre of large diameter may enter the neurilemmal sheath of a small diameter fibre and so be unable to attain its normal diameter; when connections are re-established the fibre will conduct at the lower velocity characteristic of the new diameter.

(v) The new fibres in neurilemmal tubes of any size increase in diameter very slowly, so that even months after the onset of regeneration none of the large fast conducting fibres have yet appeared; the functional properties of the nerve are consequently modified.

(vi) The rate of re-establishment of peripheral connections depends on the length of the route to be traversed and on the difficulties encountered, especially those met while crossing the original gap.

(2) FUNCTIONAL RESULTS OF IMPERFECTIONS OF REGENERATION.—Functional recovery is likewise imperfect.

(i) If many skin receptors fail to establish central connections, skin sensitivity is generally depressed and areas of anæsthesia may be present.

(ii) The ability to localize a stimulus is impaired when abnormal connections are established between the periphery and the cerebral cortex. Thus, suppose that point *a* at the base of the thumb was connected originally with area *A* in the sensory cortex; "experience" has "taught" us that cortical activity at *A* represents stimulation of point *a*. If new peripheral connections are established, cortical area *A* may become connected with, say, point *b* at the tip of the thumb. A stimulus at *b* will then be interpreted centrally as coming from *a*, *i.e.* there will be false localization which can only be overcome, if at all, by prolonged practice (cf. p. 756). Or again, cortical touch area *A* may become connected with a skin temperature receptor; temperature stimuli will consequently not be recognized at all or be misinterpreted as touch stimuli.

(iii) Because of the altered pattern of fibre diameter, the impulses which reach the cortex may also undergo modification in other ways, *e.g.* in velocity or in the character of the action potential.

The interpretation of the quality, intensity, and locality of a stimulus depends on the perfect action of the whole afferent system from periphery to cortex (as fully discussed on p. 570). The facts set out above explain why physiological recovery is always to some extent imperfect. Return of function takes a long time; the initial phases of recovery are associated with the return of the cruder aspects of sensation, with false localization, and false appreciation of quality and intensity; accurate sensation returns very slowly, partially or at best not quite completely.

(3) RECOVERY IN THE HEAD TYPE OF EXPERIMENT.—In Head's experiment, in which the anatomical conditions were almost ideal, recovery occurred in two fairly well defined stages. The first phase of crude recovery (*protopathic sensibility*) began after 8 weeks and was maximal in 30 weeks. The returning sensation was punctate, being limited to some *pain* spots, *hot* spots, and *cold* spots, the intervening areas of skin being insensitive. The findings at this stage are as follows:

(i) *Pain.*—A pin-prick cannot be located accurately, and the pain radiates widely and is not infrequently referred to some part at a distance from the point actually stimulated. A stronger pain stimulus than normal must be applied before it can be appreciated—in other words, the threshold for pain is high. The reaction excited is excessive and "sickly" in character; the

stimulus gives rise to greater pain on the affected side than does a like stimulus on the normal side. The *first* return of pain is associated anatomically with the growth of a few isolated pain fibrils into the cutaneous nerve plexuses. As might be expected, deep pricks (2–3 mm. deep) may arouse pain in areas which are still anæsthetic to lighter and more superficial pricks.

(ii) *Temperature.*—A sensation of cold is produced by temperatures below 24° C.; similarly, a sensation of heat results from temperatures above that of the body—the lower limit being between 38° and 45° C. It is impossible to recognize any temperature quality between 24° and 38° C. Thus extremes of temperature are recognized as "cold" or "hot" respectively, while intermediate grades give rise to no sensation of temperature whatever.

(iii) During this period trophic changes in the skin, which may have resulted from unappreciated cuts and burns, begin to recover rapidly. Normal growth of the part is also restored.

During the next year or longer, the finer and more discriminative aspects of sensibility (*epicritic sensibility*) return. Intermediate grades of *temperature* can now be recognized. *Tactile localization*—the ability to recognize the exact spot touched—is restored. *Tactile discrimination* returns: two compass-points applied close together are recognized as two distinct points. Light touch, *e.g.* the application of cotton wool, is now appreciated. Abnormal radiation ceases, and the quality of the sensation becomes normal in character.

The two stages described by Head often overlap to a varying extent; functional recovery is less satisfactory in most patients with nerve injuries than in his case.[1]

Dorsal Nerve Roots.[2]—All the afferent impulses from the periphery, (including the *viscera* (p. 735), pass into the dorsal *nerve roots* and enter the spinal cord. The medullated A fibres in the dorsal roots vary in diameter from 1–20 μ; in addition, there are numerous *non-medullated* C fibres present, constituting about 40% of the total number. The dorsal roots also contain *vasodilator* (probably C) fibres (p. 308).

Skin Distribution of the Dorsal Roots. Dermatomes.—The skin area supplied by a dorsal nerve root (*dermatome*) cannot be mapped out simply by determining the extent of the anæsthetic zone resulting from section of the particular root, owing to the extensive degree of *overlap* between adjacent dorsal roots. The methods which have been employed in man are:

(i) Cut *three* dorsal roots (in operations carried out for the relief of intractable pain) above and three below the root investigated and determine the area of *residual sensibility*; this area gives the full extent of the intact dermatome.

(ii) Stimulate the peripheral end of the cut dorsal root (exposed at operation) and map out the area of resulting cutaneous vasodilatation, or mark out the area of herpetic eruption in pathological irritation of the dorsal root ganglia (p. 309).

Afferent Paths in the Spinal Cord.—1. Cutaneous Sensibility.—The fibres subserving skin pain, temperature, and a part of those for touch end

[1] When a *motor* nerve is cut, similar anatomical and physiological difficulties arise. Motor fibres may grow to new muscle fibres in the same or in another muscle, or may form attachments to sensory endings. Deep sensibility may be impaired and the skill of muscle movements is reduced. When a *mixed* nerve is cut, further complications may occur, as motor fibres may grow to the skin, and skin fibres to muscles.

[2] The law of Bell-Magendie states that the dorsal roots contain exclusively afferent and the ventral roots exclusively efferent fibres.

round the cells of the dorsal horn of grey matter. The fibres which arise here cross to the ventro-lateral columns of white matter of the opposite side to constitute the *spinothalamic* tract (Fig. 355). This tract consists of two divisions: a *lateral* one, carrying pain and temperature fibres, and a *ventral* one carrying touch fibres. The spinothalamic tract also contains: pain fibres from muscles and related structures, and viscera; fibres subserving sexual and bladder sensations; fibres subserving tickle, itch, and muscular fatigue. Some of the fibres for touch as explained below, pass into the dorsal columns of the same side (cf. p. 560).

The crossing of the spinothalamic fibres takes place round the central canal of the spinal cord, in the grey commissure. In the lower levels of the spinal cord these fibres cross transversely in the segment from which they arise. In the higher levels of the cord this crossing becomes increasingly oblique. Thus, in the thoracic region, it occupies two or three segments, and in the cervical region four or five segments. At all levels the fibres for pain cross most transversely, those for touch most obliquely, while those for temperature occupy an intermediate position (cf. p. 696).

2. Deep Sensibility.—The fibres which carry impulses from the deep structures end in three ways:

(1) Some end round Clarke's column of cells at the base of the dorsal horn, chiefly in the thoracic and upper lumbar regions. From this column the dorsal spinocerebellar tract and the ventral spinocerebellar tract arise. These two tracts convey impulses from the muscles, tendons, and joints to the cerebellum; these impulses do *not* give rise to conscious sensations. The effect of an injury to these cerebellar tracts is to "maim" the cerebellum; unless the cerebellum continually receives impulses from the deep structures it cannot carry out its functions of maintaining posture and co-ordinating the activities of muscles. We therefore find that injuries to the ascending cerebellar tracts produce symptoms of cerebellar disease such as ataxy or atonia. (These are described more fully on pp. 608 *et seq.*)

(2) Some end in the spinal grey matter or pass up to the brain stem nuclei (*e.g.* vestibular or reticular nuclei, red nucleus) and are concerned with reflex posture (tone) and reflex movements.

(3) The other fibres from the deep structures pass up directly in the dorsal columns of the spinal cord. In the cervical region the dorsal columns are divided by a septum into the medial column, the *funiculus gracilis* or *tract of Goll*, and a lateral column, the *funiculus cuneatus* or *tract of Burdach.*

(i) The *funiculus gracilis* arises from ganglion cells of dorsal nerve roots supplying the lower half of the body and therefore carries impulses from the lower limb and the lower half of the trunk. Every fibre as it enters the dorsal columns of the cord comes to lie close to the medial side of the dorsal horn. As it ascends the cord it is gradually displaced medially by fibres which enter at higher levels. The *funiculus gracilis* in the lumbar region thus lies close to the medial side of the dorsal horn; as it ascends it passes medially, and in the cervical region it occupies the medial column of white matter. The tract ends in the nucleus gracilis in the medulla.

(ii) The *funiculus cuneatus* arises from ganglion cells of dorsal nerve roots supplying the upper limb and upper half of the trunk. It occupies the lateral part of the dorsal column of white matter, and ends in the nucleus cuneatus in the medulla.

A further relay of fibres from these medullary nuclei passes to the thalamus and thence to the sensory cortex. There the impulses give rise to sensations about the *position* of the limbs in space, and the direction and extent of *movements* whether passive or active. Injury to the dorsal columns results in loss of conscious sensation from the deep structures and disturbance of movements (for details see p. 561 (2) and p. 562 (4)).

The dorsal columns also contain *touch* fibres from the same side of the body; these touch fibres do not relay in the spinal cord but travel up the dorsal columns to end in the gracile and cuneate nuclei in the medulla oblongata (Fig. 355).

Sensory Disturbances in Diseases of the Spinal Cord.—(i) SYRINGOMYELIA is a condition of excessive overgrowth of neuroglial tissue, accompanied by cavity formation involving the grey matter round the central canal of the spinal cord. The crossing fibres subserving pain and temperature, which decussate in the grey commissure, are damaged, with resulting loss of these sensations. The touch fibres which cross in this region are likewise destroyed, but touch has a double path; the fibres which ascend in the dorsal columns escape. We thus obtain the characteristic *dissociated anæsthesia* of this disease, *i.e.* there is loss of pain and temperature sensibility, while the sense of touch is retained.

Hæmatomyelia is a condition in which a hæmorrhage suddenly destroys the same region of the cord as in syringomyelia; similar sensory findings are found. The same changes are found in *intramedullary* tumours of the cord.

The nutritional or *trophic* changes in skin and bone in syringomyelia may be due to loss of the protective pain and temperature sensations. Any part that is injured fails to receive care and attention because no pain is felt; suppurating whitlows and necrosis of bone are common.

(ii) In SUBACUTE COMBINED DEGENERATION of the cord (associated with pernicious anæmia (p. 197)) the *dorsal columns* are often affected early, with consequent loss of so-called conscious muscle sense.

(iii) In HEMISECTION OF THE CORD by injury or by compression from extramedullary growth, dorsal column sensibility is lost on the same side as the lesion; pain and temperature sense are lost and touch is blunted on the opposite side (cf. p. 695).

Section of Dorsal Nerve Roots.—Section of the *dorsal nerve roots* produces the following results in the corresponding segments of the body:[1]

(i) Loss of all forms of sensation—pain, temperature, touch, muscle and visceral sensibility; trophic changes appear.

(ii) Loss of all reflexes superficial and deep; loss of muscle tone.

(iii) Great clumsiness in the movement of the part, which makes its use almost impossible. This disturbance develops because all parts of the central nervous system—the cerebral cortex and the centres concerned with the reflex control of posture—are deprived of afferent impulses from the muscles. The motor impairment is well shown by comparing a cat in which all the dorsal nerve roots to the fore limbs are cut with another in which the skin alone is denervated. The second animal can readily walk up an inclined ladder, while the former can never reach the top and usually falls off the ladder

[1] For fuller details of results of clinical injuries to the dorsal nerve roots, see the discussion of tabes (p. 561).

and has to watch the rungs very carefully all the time. If, in a monkey, the dorsal nerve root supply to a limb is severed, it ceases to be used altogether even when the animal is in grave danger.

Tabes Dorsalis—The essential lesion in this common disease is a degeneration of the dorsal nerve roots central to their ganglia, affecting especially the fibres which ascend in the dorsal columns (*i.e.* those concerned with *deep sensibility*) and those which subserve pain.

The clinical manifestations illustrate well the effects of *partial* destruction of dorsal nerve roots, and amplify the brief summary given in the preceding section. The resulting phenomena are subjective and objective :

1. Subjective Phenomena.—*Lightning pains* : these come in attacks, with intervals of freedom. The first effect of the disease seems to be to stimulate pain fibres in the dorsal nerve roots. The pain may be referred to skin, muscle, or bone, and may vary in intensity from slight discomfort to intolerable agony.

2. Objective Findings.—The fibres conveying *pain* impulses and those from the *deep structures* are affected first and to the greatest extent.

(1) *Pain Sensibility.*—Pain sensibility may be lost, its appreciation is delayed (p. 554), or it may be inaccurately localized. As in syringomyelia, loss of pain sensibility may be responsible for *trophic* disturbances. The perforating ulcers found under the ball of the foot may arise from the neglect of a corn.

As the dorsal roots mainly affected in tabes are those which supply the lumbo-sacral and the cervico-thoracic regions of the cord, the common areas of anæsthesia are round the anus, over the legs, upper chest, and the ulnar borders of the hands. The involvement of the fifth nerve accounts for the anæsthesia of the central part of the face.

The joint deformity known as the *Charcot joint* may be produced thus : an injury is inflicted on a joint, or a mild sprain or subluxation may occur ; as no pain is aroused, the condition is neglected, receives no effective treatment, and becomes progressively worse until considerable damage to the articular surfaces of the bones results.

(2) *Deep sensibility.*—As the fibres that ascend the *dorsal columns* are damaged there is loss of sense of position, passive movement, and vibration sense in bone.

(i) Loss of *sense of position* : with the eyes closed the patient is unaware of the exact position of the various parts of his body.

(ii) Loss of *sense of passive movement* : while the eyes are closed, the great toe, for example, may be flexed or extended without the patient's knowledge of the alteration of position imposed on the toe.

(iii) Loss of *vibration sense in bone* : after striking a tuning-fork, the handle is placed on some bony prominence. Normally, a sensation is felt which is varyingly described as a " buzzing " or " electrical " sensation. When dorsal column sensibility is lost, the patient states that he merely feels a cold object in contact with his skin.

(3) *Reflexes.*—*Muscle tone* depends on impulses which enter via the dorsal nerve roots from the corresponding muscles (p. 582). As these impulses are not received, loss of tone (atonia) or diminished tone (hypotonia) results. The muscles are flabby ; the limbs can be placed in unnatural positions, and excessive movement is permitted at the joints without producing discomfort.

The *deep reflexes* (p. 644) (*e.g.* knee-jerk, ankle-jerk) which depend on the integrity of "stretch afferents" in the dorsal nerve roots, are also lost.

(4) *Voluntary Movement.*—There is considerable disturbance of *voluntary movement,* illustrating the important general principle that the indispensable basis of all purposeful and effective motor activity is accurate information. Normal persons carry out many movements without the aid of the visual sense, being guided partly by the impressions which reach the sensory cortex via the dorsal columns, and partly by proprioceptive impulses which do not attain consciousness but end in the spinal cord, brain stem, and cerebellum to control posture and indirectly to guide voluntary movement (p. 606). Many movements, *e.g.* walking, can normally be carried out with a minimum of attention. When the afferent impulses from muscles are cut off, the gait becomes clumsy; the patient walks with his legs apart to help compensate for his ataxy. Every phase of each limb movement is overdone, because the normal afferent impulses which would check the movement are not received. The feet are raised too high and stamped down too forcibly—*stamping gait.* The *finger-nose test* demonstrates the defect under consideration very well. The patient is asked to place the tip of the forefinger of the outstretched hand on the nose while his eyes are shut. Instead of taking the shortest and most direct route, and readily touching the nose, the finger wanders about aimlessly, and finally alights on the cheek, chin, or neck. As the patient is unaware of the position of the nose or the finger and does not know what progress the hand is making, it is little wonder that he fails to find his objective. Should the patient close his eyes while standing with the feet together, he sways from side to side, and, if not supported, finally falls to the ground (*Romberg's sign*). This inability to stand may be due in part to loss of the afferent impulses which subserve the stretch reflexes; these normally produce a sustained contraction of the antigravity muscles in the legs and trunk thus maintaining the erect posture (cf. p. 595).

(5) *Temperature* and *touch sensibility* are affected to a variable extent, but for some unknown reason, the dorsal root fibres subserving these sensations tend to escape injury.

(6) The *sphincter* trouble in tabes results from the loss of visceral sensibility of bladder and rectum. The patient does not know when his bladder is comfortably full, and consequently allows it to go on filling, and become excessively distended without experiencing any desire to micturate. The heightened pressure in the bladder finally forces the sphincter open, and small amounts of urine overflow at intervals (see also p. 771).

Ascending Paths in the Brain Stem.—(1) FUNICULUS GRACILIS AND CUNEATUS.—In the lower medulla, the funiculus gracilis ends in the nucleus gracilis, the funiculus cuneatus in the nucleus cuneatus. From the small cells forming these nuclei a fresh relay of fibres arises, which sweeps ventrally across the middle line as the internal arcuate fibres to form the *medial lemniscus* which lies near the middle line and dorsal to the pyramid. The decussation is called the decussation of the medial lemnisci. Some fibres from the nuclei pass via the external arcuate fibres and the restiform bodies (inferior peducles) to the palæocerebellum (Fig. 383A and p. 606).

The medial lemniscus at this stage is only conveying the impulses carried in the dorsal columns, *i.e.* those from receptors in muscles and tendons (deep sensibility) and from touch receptors in skin.

(2) SPINOTHALAMIC TRACT.—The fibres which carry impulses from pain and temperature receptors (p. 559) remain on the surface of the medulla lateral to the olivary nucleus, where they mingle with the ventral spino-cerebellar tract. The fibres from touch receptors diverge medially from these to rejoin the rest of the touch fibres (which are lying in the medial lemniscus). In the pons, the pain and temperature fibres of the spinothalamic tract have also passed medially to join the main sensory path. The medial lemniscus thus enlarged, is now carrying all the impulses from the skin and muscles of the opposite side of the body below the head (Fig. 355).

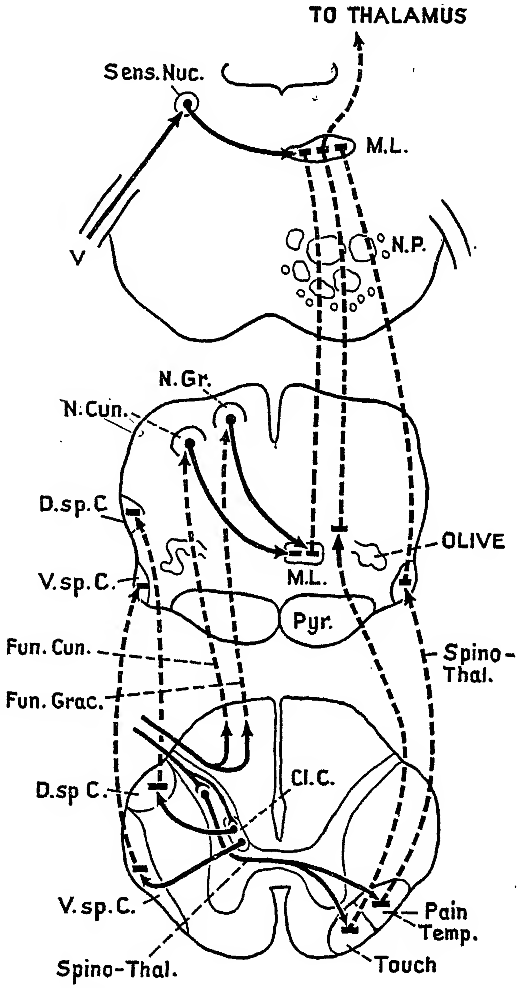


FIG. 355.—Diagram of Course of Ascending Tracts.

V=Afferent fibres of fifth nerve. Sens.Nuc.=Sensory nucleus of fifth nerve. N.P.=Nuclei pontis. Pyr.=Pryamidal tract. N.Gr.=Nucleus gracilis. N.Cun.=Nucleus cuneatus. D.Sp.C. =Dorsal spinocerebellar tract. V.Sp.C.=Ventral spino-cerebellar tract. M.L.=Medial lemniscus. Ol.=Olivary nucleus. Spino-Thal.=Spinothalamic tract. Fun.Cun., Fun. Grac.=Funiculus cuneatus and gracilis. Cl.C.=Clarke's column of cells.

(3) CRANIAL NUCLEI.—Afferent impulses from the *face* enter in the fifth nerve, and end in the principal sensory nucleus in the pons, and in the long descending root and adjacent grey matter which extend down to the upper cervical cord. From this long column of grey matter a fresh relay arises which crosses the middle line to enter the medial lemniscus. The fibres of the ophthalmic division of the fifth nerve end in the lowest part of the spinal part of the descending root; the maxillary division ends in the middle part of the root, while the superior part of the nucleus receives the inferior or mandibular division of the nerve. In other words, the face is represented "upside-down" in the nucleus. The medial lemniscus also receives accessions of fibres from other *sensory cranial nuclei*, particularly

the vagus (from the air-passages), the dorsal nuclei of the ninth and seventh nerves (*taste* fibres) (for details see p. 581) and perhaps from the vestibular division of the eighth nerve (from the labyrinth) (p. 597).

In the midbrain the medial lemniscus lies in the tegmentum dorsal to the substantia nigra and close to the middle line; it passes through the subthalamic region to end in the thalamus (Fig. 355A).

The Thalamus[1] (Fig. 356).—The main nuclei and connections of the thalamus are set out below in summary form:

1. Anterior Nucleus (Fig. 356, 1, *A*).
   (i) Receives the mammillo-thalamic tract which relays fibres that arise in the hippocampus and travel via the fornix to the mammillary bodies (cf. p. 582).
   (ii) Sends fibres to the cingular gyrus (area 24) and the paracentral lobule.

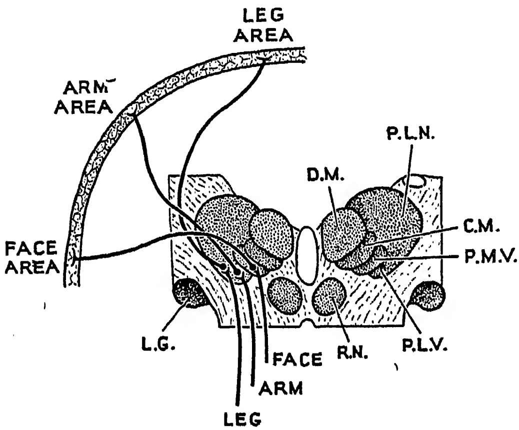


Fig. 355A.—Ascending Tracts to Thalamus and Sensory Cortex. (See legend to Fig. 356.)

2. Medial (Dorso-medial) Nucleus (and related "intralaminar" nuclei) (Fig. 356, 1, *M*; 2, *DM*; 3, *DM*, *CM*):
   (i) Receives and sends fibres to the hypothalamus (Fig. 428).
   (ii) Sends fibres to most regions of the frontal lobes (p. 669).

As explained on p. 670 rich to-and-fro fibre connections link up the anterior and medial thalamic nuclei with the hypothalamus and the prefrontal lobes, and so convert these grey masses into an integrated functioning entity.

In addition these two thalamic nuclei establish to-and-fro connections with most areas of the cerebral cortex and thus subserve the "resting" electroencephalogram (p. 617; Fig. 357). The suppressor bands in the cerebral cortex send fibres via the caudate nucleus to the thalamus to block this "closed cortico-thalamic circuit" and so suppress the cortical "resting" potentials (p. 620).

3. Lateral Nucleus.—This is essentially a relay station, linking skin, muscle, taste receptors and efferent fibres from the cerebellum with the postcentral and precentral cortex.
   (1) *Lateral part*: (i) dorsal (Fig. 356, 2, 3, *DL*); (ii) posterior (Fig. 356, 3, *PL*); these send fibres to the parietal lobule.
   (2) *Ventral part*:
      (i) *Anterior* (Fig. 356, 1, *AV*): sends fibres to the caudate nucleus and putamen.
      (ii) *Lateral* (*latero-ventral nucleus*) (Fig. 356, 2, *LV*): receives fibres

[1] Walker, *Primate Thalamus*, Chicago, 1938.

from the opposite neocerebellum via the crossing superior cerebellar peduncle; relays the cerebellar impulses to the excitomotor areas 4 and 6.

(iii) *Posterior* (*postero-ventral nucleus*):

(*a*) *medial* (Fig. 356, 3, *PMV*): receives the fibres of the secondary neurones from the face (V) and the taste receptors in the tongue (via IX and VII); relays the impulses to the

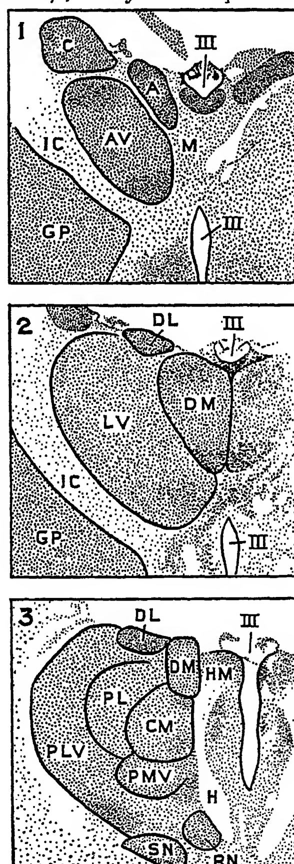


FIG. 356.—Nuclei of Thalamus (Monkey) (after Ranson & Clark, *Anatomy of Nervous System, Phila.*, 8th edn., 1947.)

The sections are in the frontal plane; 1, 2, 3 are in that order from before backwards.

A, anterior thalamic nucleus.
AV, anterior ventral nucleus.
C, caudate nucleus.
CM, centrum medianum.
DL, dorsal lateral nucleus.
DM, dorsal medial nucleus.
GP, globus pallidus.
HM, medial habenular nucleus.
H, habenulopeduncular tract.
IC, internal capsule.
LV, lateral ventral nucleus.
M, nuclei of midline.
PL, posterior lateral nucleus.
PLV, posterolateral ventral nucleus.
PMV, posteromedial ventral nucleus.
RN, red nucleus.
SN, substantia nigra.
III, 3rd ventricle.

face area and the associated taste area in the inferior part of the postcentral cortex.

(*b*) *lateral* (Fig. 365, 3, *PLV*): receives the fibres of the secondary neurones from the arm and leg; relays the impulses to the middle and superior part of the postcentral cortex (arm and leg areas).

4. POSTERIOR NUCLEUS:

(1) *Pulvinar*: sends fibres to the parietal "association areas."

(2) *Medial* and *lateral geniculate bodies*: receive the fibres of the

secondary visual and auditory neurones and relay the impulses to the visual and auditory areas of the cortex (pp. 572, 575).

Functions of the Thalamus.—(1) The thalamus (medial nucleus), by virtue of its efferent connections with the hypothalamus is a reflex centre concerned with *emotional exteriorization* (for full discussion see p. 664).

(2) Because of its rich interconnections with the prefrontal lobes the thalamus is part of an integrated nervous complex subserving *personality*, *emotional affects*, and *social behaviour* (p. 670).

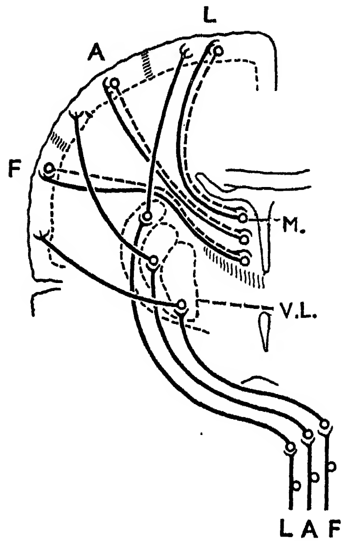


FIG. 357.—Connections of Postero-ventral Nucleus and Medial Nucleus of Thalamus.

L, A F=Afferent fibres from leg, arm and face which relay in the brain stem and cross over to the opposite side in medial lemniscus to end in the posteroventral nucleus of thalamus (V.L.). They relay there to pass to the leg, arm and face areas of the sensory cortex.

M=The medial thalamus, which is connected with the cortex by a "private" thalamo-cortical and corticothalamic pathway. This to-and-fro path is the circuit which is responsible for the resting electro-encephalogram.

(3) The thalamus is concerned with the maintenance and control of the resting *electroencephalogram* (p. 617; Fig. 357).

These intimate thalamo-cortical inter-relationships make the anterior and medial nuclei of the thalamus (for most purposes) an integral, though anatomically displaced, part of the cerebral cortex itself.

(4) The interconnections of the thalamus and *corpus striatum* are shown in Fig. 424; the relationship of the thalamus to the syndrome of choreoathetosis is considered on p. 661.

(5) The latero-ventral nucleus relays *cerebellar* impulses to the excitomotor areas. Lesions of this nucleus produce signs of cerebellar disturbance in the opposite side of the body. These consist of weakness and slowness of movement; ataxy during voluntary movement (cf. p. 610); diminished muscle tone. The superficial reflexes and the plantar response are normal on the affected side of the body.

(6) The postero-ventral nucleus relays *skin, muscle,* and *taste* afferents to the postcentral cortex. Lesions of this nucleus produce striking sensory disturbances (see p. 567).

(7) RELATION OF THE THALAMUS TO CONSCIOUS SENSATION.—The following observations are relevant to the discussion:

(i) Removal of the cerebral cortex on one side in the monkey or baboon produces on the opposite side of the body complete anæsthesia for a few days; subsequently limited recovery occurs. If the other cerebral cortex is now removed this residual sensory activity remains unimpaired, indicating that it is mediated by a *subcortical* level.

(ii) In man, hemidecortication produces severe sensory loss over the opposite trunk and limbs; all forms of sensation are lost except heavy touch and pin-prick which are recognized, but imperfectly localized. Over the opposite *face* the sensory disturbance is less marked: touch and pin-prick are readily recognized and to some extent localized. There is evidence,

however, that the greater preservation of skin sensibility in the face is due to the fact that the skin of the head is represented in both hemispheres. As some sensibility persists in the trunk and limbs after hemidecortication it seems likely that in man subcortical neurones, probably in the thalamus, can mediate some form of " vague " sensation.

(iii) *Thalamic Syndrome.*—This clinical syndrome is due to occlusion of the artery which supplies (mainly) the portions of the lateral thalamic nucleus, which are the relay station for the arm, leg, and cerebellar afferents. All the symptoms are found on the opposite side of the body and are as follows :

(*a*) Weakness, decreased muscle tone, and ataxy, due to injury to the cerebellar afferents (pp. 608 *et seq.*).

(*b*) Choreic and athetotic movements (p. 660), due perhaps to interference with the connections between the thalamus and the corpus striatum.

(*c*) The same kind of loss of the discriminative aspects of sensation as

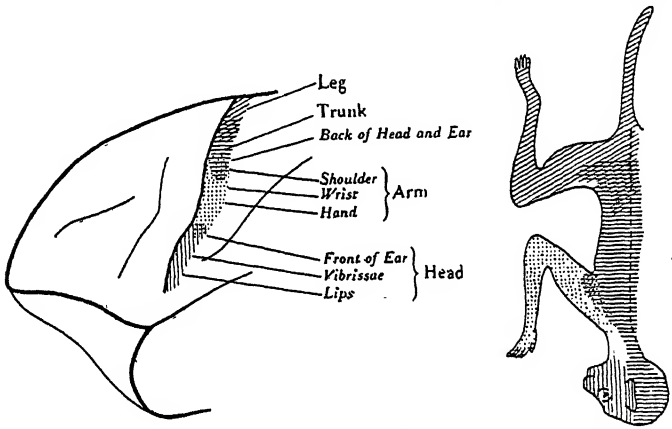


FIG. 358.—Projection of Body Wall of Monkey on Sensory Cortex. (Adrian, *J. Physiol.*, 1941, *100*.)

occurs in lesions of the postcentral cortex. The loss includes the sensation of light touch, tactile localization and discrimination, intermediate grades of temperature and appreciation of small movements at joints (cf. p. 569). These results are readily understandable as the thalamic nucleus which relays to the cortex has been damaged. The sensations which are retained are appreciation of gross movements and of hot and cold stimuli (which are poorly localized).

(*d*) There are *altered emotional affects.* Thus one of Head's patients complained that his trousers produced such unpleasant feelings that he had to remove them. Another patient found the singing of hymns and of " God save the King " quite intolerable. Some stimuli, *e.g.* warmth applied to the skin, tickling, and erotic stimuli may produce greatly enhanced states of pleasure ; on the other hand painful stimuli (*e.g.* pin-prick) may produce intolerable anguish. " Spontaneous " agonizing pain may develop without obvious cause.

There is no satisfactory explanation of these emotional disturbances.

Head attributed them to the release of the anterior and medial thalamic nuclei from cortical inhibitory control; but the lesion does not usually involve corticothalamic fibres. Alternatively it may be suggested that the anterior and medial nuclei are released from some other *thalamic* influence; as a result the grey " complex " consisting of anterior and medial thalamus and frontal lobes, functions abnormally producing the changed " affects " described.

Somatic Sensory Cortex.[1]—The functions of that part of the parietal cortex which subserves conscious skin and muscle sensibility have been studied in a number of ways.

(1) Cortical Potentials in Animals.—The points of immediate termination of afferent impulses in the sensory areas (receiving areas, " arrival platforms ") of the cortex can easily be mapped out by electrical means. The cortex is exposed in an anæsthetized animal; stimulation of skin or muscle sets up a burst of action potentials in a restricted region of the cortex which represents the arrival point of the impulses. The map thus determined when the body wall (skin and muscles) is stimulated in the monkey is shown in Fig. 358. The receiving area in all species is mainly concerned with those

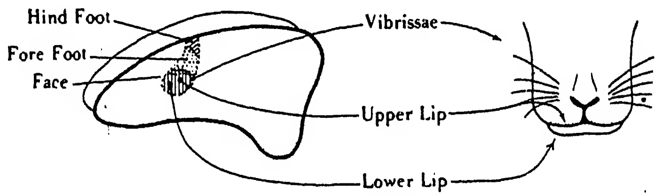


Fig. 359.—Representation of Different Regions of Body Wall on Sensory Cortex in Rabbit

The "face" area represents the lips and vibrissæ exclusively. (Adrian, *J. Physiol.*, 1941, *100*.)

parts of the body which are mostly closely related to the outside world; thus in the cat there is a large area for the claws, in the rabbit for the mouth region (Fig. 359), in the dog for the face, and in the monkey for the hands and face. The impulse patterns set up by the peripheral receptors become modified as they pass through the various relays in the central nervous system as a result of central summation, irradiation, and after-discharge; the patterns arriving in the sensory area thus differ significantly from those set up by the sensory endings. Strangely enough, no cortical responses have yet been detected in the animal species examined, following the application of temperature or painful stimuli.

(2) Anatomical Studies.—Such studies in man show that the fibres of the medial lemniscus are topographically projected[2] on to the postcentral (ascending parietal) gyrus. The postcentral gyrus has the typical granular structure which characterizes the sensory cortical areas (Fig. 388; p. 615). Fibres from other parts of the thalamus (dorsal-lateral, posterior-lateral nuclei and pulvinar which receive *no* incoming sensory fibres) project to the

[1] Bard, *Harvey Lectures*, 1937-38, *33*, 143. Bard *et al.*, *Johns Hopkins Hosp. Bull.*, 1942, *70*, 399.

[2] Topographical projection: this means that each small group of arriving fibres makes connection with a limited group of cells found in a specific cortical area.

parietal lobe *behind* the postcentral gyrus. This latter region therefore receives what has been called evasively the "integrated product of thalamic activity."

(3) Electrical Stimulation in Man.—Electrical stimulation of the exposed surface of the postcentral gyrus in a conscious man leads to hallucinations of tactile stimulation, feelings of numbness or tingling or a sense of movement or pressure which are referred to the opposite side of the body; but there is *never any complaint of pain.* Irritation arising from disease produces similar symptoms. Stronger electrical stimulation may set up movements owing to impulses passing forward to the motor areas. A high degree of topographical representation of the different parts of the body can be demonstrated by electrical methods in the postcentral gyrus in man corresponding roughly to the motor localization in the precentral cortex (cf. p. 635), though more overlapping takes place.

(4) Clinical Evidence.—The postcentral gyrus, which directly receives the afferent paths, is presumably mainly concerned with mediating the elementary sensations of touch, pressure, heat, cold. The more posterior part of the parietal lobe (superior parietal lobule, part of the supramarginal and angular gyri) is thought to be associated with the more elaborate processes of *discriminating between stimuli* and the *recognition of common objects placed in the hand* (*stereognosis*) without looking at them. Study of patients with disease or injury to the parietal lobe shows that the parietal cortex is concerned with the following aspects of sensation:

(i) Differences in the relative *intensity* of different stimuli: heat is not merely distinguished from cold, but warm objects are distinguished from warmer, cold from colder, rough from rougher, and so forth.

(ii) Recognition of *spatial relationships*: (*a*) *Tactile localization*: the precise point stimulated is accurately located. (*b*) *Tactile* (*two-point*) *discrimination*: two points of a compass placed close together are recognized as two and not as one. (*c*) The extent and direction of *small joint displacements* can be estimated accurately. Thus the relations of a stimulus in one, two, or three dimensional space are clearly defined.

(iii) Appreciation of *similarity and difference* in external objects brought in contact with the surface of the body (without the aid of visual impressions): differences and similarity of size, weight, form, and texture are thus recognized.

(iv) *Stereognosis*: this is the most elaborate function subserved by the parietal cortex. It necessitates perfect reception of the impulses set up by the stimuli from the object. The sensations produced are "synthesized" in the cortex and compared with previous similar sensory "memories."[1] We thus recognize an object as a half-crown and distinguish it from a penny piece by the presence of the milled edge (for which we deliberately feel).

It should be emphasized that the forms of sensation enumerated in paragraphs (i)–(iv) above are lost or impaired in lesions of the parietal lobe. (As explained on p. 581 the face area of the postcentral gyrus is closely associated with the taste area.)

As gross forms of anæsthesia are rare in clinical lesions of the sensory cortex casual examination of the sensory nervous system may reveal no abnormality. The patient must always be tested for stereognosis (which is almost constantly defective) and in the other ways indicated above. Prolonged examination at one time must be avoided, as the sensory cortex fatigues readily.

[1] **The words in inverted commas are, physiologically, meaningless.**

As explained on p. 566, hemidecortication produces more grave sensory loss; the residual sensibility in such cases is due partly to activity of the sensory cortex on the opposite side and partly to activity of the thalamus itself.

Motor Changes in Lesions of Sensory Cortex.—In lesions of the sensory cortex the patient's complaint is commonly of clumsiness (ataxy) and "uselessness" of the affected side (which is on the side opposite to that of the diseased cortex). This finding again emphasizes the overwhelming importance of proper afferent guidance of the motor areas of the brain.

General Discussion of Relation of Stimulus to Sensation.—(1) PROJECTION AND LOCALIZATION.—Activity of the sensory cortex (whether set up naturally or artificially) is not experienced in consciousness as an event occurring in the head but as an event at the periphery; this *projection* of the sensation varies with the receptors involved. Thus impulses from the body wall give rise to sensations projected to skin or muscle; from the labyrinth—to the head and neck; from certain viscera—somewhat vaguely to the interior of the body; from the *distance receptors* (eyes, ears)—to external space whence the stimuli originated (thus we project retinal impulses set up by a man seated in an armchair to the place where he is and not to our own eyes; we project a sound to the orchestra whence it came, and not to our own ears). A special aspect of projection is *localization* of stimuli applied *e.g.* to skin or muscle. Localization depends primarily on *regular anatomical connections* between the receptive field (*e.g.* skin) and the receiving area of the cortex, *i.e.* on *topographical* (so-called "*point-to-point*") projection of the peripheral field on to the receiving cortical area. The electrical evidence mentioned on p. 568 supports this view. For highly accurate localization, a prerequisite is that minute areas of the sensory surface should connect with "private" groups of cortical cells; this is the case with the skin of the fingers. On the back, however, large skin areas are projected on to a small cortical zone; precision of localization is correspondingly poorer. The term "point-to-point" connection must not be understood too literally. An area of 1 sq. mm. of human skin may contain ten groups of two or three touch corpuscles supplied by nerve fibres which enter several dorsal root filaments or even two separate roots. Consider further, that the touch fibres have a dual central path (dorsal columns and spinothalamic tract) and that at each relay each afferent fibre establishes multiple potential connections. It follows that any "spot" on the skin probably activates not a cortical spot but a *zone*: the central region of the zone is stimulated maximally, the more peripheral parts feebly; the cortical response to a punctate stimulus is thus a characteristically *patterned* localized area of activity, which in some unknown way gets attached to it spatial properties with respect to the sensory surface. When a more extensive skin surface is stimulated, multiple cortical areas are activated. (For *referred pain* see p. 756).

(2) INTENSITY of the sensation is related to the frequency of the afferent impulses (p. 551).

(3) QUALITY (MODALITY) OF THE SENSATION.—This seems to depend on unknown special properties of the region of the cortex activated. No matter how a sensory cortical "centre" is stirred up to activity—whether by natural stimulation of the appropriate nerve endings, or by artificial stimulation of the afferent nerve trunks, or of the cortical substance itself—the response

always has the distinctive and specific quality. This statement must not, however, be taken too literally; the effects in consciousness of artificial stimulation of a sensory cortical area resemble, of course, only vaguely the results of normal stimulation by impulse streams from the normal afferent channels, yet the same quality of sensation is aroused.

(4) OUTSTANDING PROBLEMS OF PAIN SENSATION.—(i) There is little evidence that the parietal cortex mediates pain sensation. The pain afferents undoubtedly pass up as far as the thalamus; lesions below the level of the thalamus may abolish pain sensibility. It seems not improbable that the complex mass "prefrontal cortex—thalamus" is responsible both for recognition of pain and the associated emotional "affect." Supporting evidence is as follows: (*a*) lesions of the thalamus may lead to "central pain"

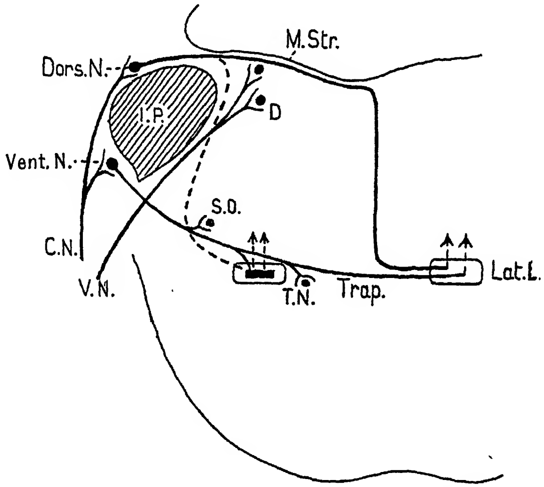


FIG. 360.—Diagram of Auditory Path.

C.N. = Cochlear Nerve; V.N. = Vestibular nerve; Dors.N. = Dorsal cochlear nucleus; Vent.N. = Ventral cochlear nucleus; M.Str. = Medullary striæ; S.O. = Superior olive; T.N. = Trapezoid nucleus; Trap. = Trapezium; Lat.L. = Lateral lemniscus; I.P. = Inferior peduncle (Restiform Body); D = Deiters' nucleus.

and greatly exaggerated pain affects (cf. *thalamic syndrome* p. 567); (*b*) in cases of visceral disease with intractable pain which has not responded to any form of drug treatment, prefrontal leucotomy may greatly decrease the "affect" ("feeling" of pain); the patient recognizes the existence of pain but it has ceased to worry him or dominate his whole outlook.

(ii) Any proffered explanation of the accurate localization of many forms of painful skin stimulation has to dispose of the following difficulties: (*a*) because of the wide interconnections of the nerve nets in the skin subserving pain any punctate skin stimulus must feed back into *multiple* afferent channels; (*b*) the thalamus can only mediate vague and inaccurate localization; (*c*) the skin is not topographically projected on to the prefrontal cortex. It will be remembered, however, that many specialized receptors have an associated accessory "pain" fibre; thus, vigorous scratching would stimulate

first touch and light-pressure receptors and then the associated pain receptors. The stimulus to touch and light-pressure receptors is accurately localized; this localization may then be "appended" to the associated pain sensation and to the "affect."

The Auditory Path.—The cell bodies of the cochlear division of the auditory nerve are in the spiral ganglion of the cochlea [1] in the internal ear; the peripheral axons pass to (or more precisely arise between) the hair cells of the organ of Corti; the central axons pass to the medulla dorsal to the restiform body. Here they end in two nuclei: the *ventral cochlear nucleus*, between the two divisions of the nerve, and the *dorsal cochlear nucleus*, which lies on the dorsolateral aspect of the restiform body. From these cell stations a fresh relay of fibres arises (Figs. 360, 441–443).

(1) From the *ventral* nucleus a band of transversely running fibres, the *trapezium*, arises, which mainly crosses to the opposite side about midway between the ventral and dorsal aspects of the pons; collaterals are given to the superior olive and the trapezoid nucleus which connect up with motor cranial nuclei (especially VI) supplying the muscles of the eyes, head and neck.[2]

(2) From the *dorsal* nucleus the *medullary striæ* [3] arise, which mainly cross the midline in the floor of the fourth ventricle and then turn sharply ventrally in the reticular formation to join the trapezium.[4] The united fibres then turn upwards as the *lateral lemniscus*.

As indicated, the decussation of the auditory fibres is partial, like that of the optic nerve fibres (*q.v.*). Some of the fibres of the trapezium and the medullary striæ pass up the lateral lemniscus of their *own* side.

In the midbrain the lateral lemniscus turns sharply dorsally to end in the *medial geniculate bodies*; the two medial geniculates are united by fibres running through the optic chiasma, forming van Gudden's commissure. Collaterals are given to the *inferior colliculi* which serve as a centrê for *reflexes* in response to sound.

From the medial geniculates a fresh relay of fibres arises, which passes as the auditory radiation behind the posterior limb of the internal capsule to end in the *auditory centre* in the temporal lobe, mainly in Heschl's gyrus (anterior transverse temporal gyrus) which is buried in the fissure of Sylvius, and also in the superior temporal gyrus. Stimulation of this region in a conscious man leads to auditory hallucinations, *e.g.* buzzing, booming, or clicking noises; animals so treated will prick up their ears. It should be emphasized that fibres from one cochlea ultimately reach the auditory cortex on both sides.

Round the primary auditory centre there is believed to be an *auditory-psychic centre* which subserves the appreciation of the nature of auditory stimuli. Thus, in the course of time, when we hear the sound of a whistle we associate it at once in our mind with a policeman, or a train moving out of a station, etc. When the significance of common sounds can no longer be appreciated, a condition of *auditory agnosia* is said to exist. In the first and second [superior and middle] temporal convolutions on the left side is an area in which the sounds employed in speech are "understood"; this constitutes the *auditory speech centre* (p. 653).

[1] Cochlea=snail-shell, or a spiral staircase, as in the Hippodrome at Byzantium.

[2] There is evidence that all the cochlear fibres relay in the superior olive and trapezoid nucleus; this would make the lateral lemniscus a tertiary auditory neurone.

[3] Or *Striæ acousticæ*.

[4] Or *trapezoid body*.

Mechanism of Hearing.—(1) Conduction of Sound Waves in Ear.—Sound waves pass down the external auditory meatus and impinge on the *drum* which, being aperiodic, reproduces faithfully the frequency of the stimulating tone. The *ossicles* of the middle ear in their turn accurately transmit the frequency of the vibrations of the drum to the footpiece of the stapes which fits into the foramen ovale. The *fluid* (perilymph and endolymph) in the *scalæ* of the internal ear in its turn vibrates with the same frequency. Owing to the anatomical arrangement of the ossicular system the energy of the vibration of the drum is transmitted to the fluid in the scalæ with a *tenfold* increase in intensity; the intensity of the vibrations of the fluid is directly proportional to the amplitude (loudness) of the tone. The frequency and energy of the vibrations of the perilymph and endolymph are transmitted to the fibres of the *basilar membrane* (*auditory strings*) on which rests the *organ of Corti*, innervated by nerve filaments from the cochlear division.

(2) Stimulation of Organ of Corti.—The fibres of the basilar membrane number 24,000 in man; they are shortest at the base of the cochlea and largest at the apex. The range of fibre length is between 64–128μ and 352–480μ. These fibres act as *resonators*; fibres of one (or closely related) length respond specifically to vibrations of a certain (probably narrow) range of frequency. The *longest* fibres respond to the *lowest* frequencies (lowest tones), the *shortest* fibres to the *highest* frequencies (highest tones). The vibrations of any section of the basilar membrane set up movements in the superjacent part of the organ of Corti; these movements serve, it is thought, as the specific stimulus to the local nerve filaments of the cochlear nerve. All the disturbances in the ear discussed so far are mechanical in nature, reproducing faithfully as stated the wave frequency and wave form of the original tone; the final result, however, in the organ of Corti is the generation of *nerve impulses*. The nerve terminals here display the same general properties as receptors elsewhere (cf. p. 549).

(i) *Analysis of Pitch.*—The pitch (*i.e.* the frequency) of the tone determines (as explained) which group of fibres of the basilar membrane responds and consequently the *locality* in the cochlea of the nerve terminals that are stimulated. Direct study of the action potentials in single or small groups of fibres in the cochlear nerve reveals no difference in the form of the potential changes in different fibres of the nerve; the cochlear action potentials furthermore are indistinguishable from those recorded from other afferent nerves. Direct experiment also proves that individual tones stimulate *specific* small groups of cochlear fibres, depending on the frequency of the tone; thus *high* tones generate impulses in the nerve fibres from the *basal* region, *low* tones in the fibres from the *apex* of the cochlea. Conversely, minute lesions of the organ of Corti result in loss of action potential changes in the (whole) cochlear nerve in response to specific tones. Thus a lesion (in the cat) 2 mm. from the basal end of the cochlea resulted in a maximal loss of response to tones of 8200 cycles; a lesion 14·5 mm. away caused a maximal loss to tones of 256 cycles; in this species about 2·5 mm. of cochlea are alloted to each octave. When a larger cochlear area is damaged, the tone loss is correspondingly more extensive.[1]

(ii) *Analysis of Loudness* (*Intensity*).—The *frequency* of the cochlear nerve

[1] Ades and Felder, *J. Neurophysiol.*, 1942, *5*, 49; Galambos and Davis, *ibid.*, 1943, *6*, 39.

impulses bears no relationship whatsoever to the frequency of the stimulating sound waves. It may be emphasized that the same lack of relationship obtains between the frequency of light waves impinging on the retina and the impulse frequency in the optic nerve, or between the frequency of heat waves falling on the skin and the impulse frequency in the corresponding cutaneous nerves. In the cochlear fibres, as in all the receptors so far examined, the impulse frequency varies directly with the *strength* of the stimulus, *i.e.* the loudness of the sound, and is thus related to the amplitude of the stimulating sound waves and the resulting force of movement of the basilar membrane. A weak tone sets up a low impulse frequency, a loud tone a high frequency. The nerve endings show little adaptation (cf. p. 550), *i.e.* the impulse frequency is not diminished when the stimulus is maintained.

(3) Auditory Centres.—The nerve impulses thus set up in the cochlea pass through the relay stations enumerated above (p. 572) to reach the cerebral cortex. The *auditory centre, i.e.* the receiving area for cochlear nerve impulses, can be readily mapped out experimentally by recording the brief action potentials which are set up in the cerebral cortex in response to sound stimuli (the same method is employed for mapping out the somatic sensory cortex (p. 568). In the monkey the centre thus delineated lies in a portion of the superior face of the superior temporal gyrus in the angle between the posterior and medial borders of the gyrus. As might be expected (cf. visual pathway, p. 576) there is an accurate topographical (point-to-point) representation of the cochlea first in the medial geniculate body and secondly in the auditory cortex. In the cat, for example, the basal end of the cochlea is projected on to the anterior end, and the apex, on the posterior end of the auditory centre, and intermediate cochlear regions intermediately on the centre. The central representation of each cochlea is *bilateral.* Corresponding points from the two cochleæ (*i.e.* those responding to the same tones) are projected on to the same spot in each auditory centre and give equivalent responses.

The fundamental mystery about the auditory centre is the same as that for all the cortical sensory areas, namely, how is it that activity of nerve cells in this region is associated with the conscious sensation of sound, and how is it that the activity of special clumps of " auditory " cells gives rise to conscious sensations of different sorts of sounds ? There is no answer as yet to these queries.

Clinical Lesions of Auditory Pathway.—Ablation (or more partial lesions) of one temporal lobe in man does not cause complete deafness but affects the acuity of hearing, localization of sound, and auditory memories; as explained, fibres from corresponding points on the cochlea end in the temporal lobe on both sides. Deafness from a lesion of the brain stem is rare, and can only occur when a lesion of the central part of the pons destroys the trapezium or involves both lateral lemnisci.

It should be mentioned for completeness that the *labyrinthine* fibres probably also terminate in the temporal lobe.

The Visual Path.—The visual fibres arise in the layer of nerve cells in the retina and pass backwards along the optic nerve to the optic chiasma. Partial decussation takes place; the fibres from the nasal sides of the retinæ cross, and those from the temporal side of the retinæ remain uncrossed. The left optic tract therefore conveys fibres from the left halves of both retinæ. Each half of the retina receives light rays from the opposite half of the field

of vision. The left optic tract corresponds, therefore, to the right or opposite half of the field of vision. The fibres from the macula lutea (yellow spot), or region of most precise vision, behave in exactly the same way. The fibres from the nasal sides of both maculæ cross, and those from the temporal sides remain uncrossed (Fig. 361).

The optic tracts thus constituted wind round the outer side of the crura cerebri and end in two main areas:

(i) The *superior colliculi* (or in the near-by *pretectal area*), which are not concerned with conscious vision but serve as a centre for visual reflexes (*e.g.* the light reflex, p. 578). Stimulation of the superior colliculi usually causes the eyes to move to the opposite side accompanied by turning of the head in the same direction, elevation of the eyebrows, opening of the palpebral fissures, and changes in pupil diameter (constriction or dilatation).

(ii) In the *lateral geniculate body*, from which a fresh relay arises which passes back in the optic radiation to the *occipital cortex.* The fibres pass through the internal capsule behind those for "common sensation," and then pass deep in the substance of the temporal lobe round the outer surface of the lateral ventricle to reach the "half-vision centre" in the occipital lobe; in man this centre is situated in the cuneus and lingual gyrus above and below the calcarine fissure on the medial aspect of the lobe (area 17) (Fig. 365). The "half-centre" is so called because, like the optic tract, each occipital centre represents the opposite half of the field of vision.

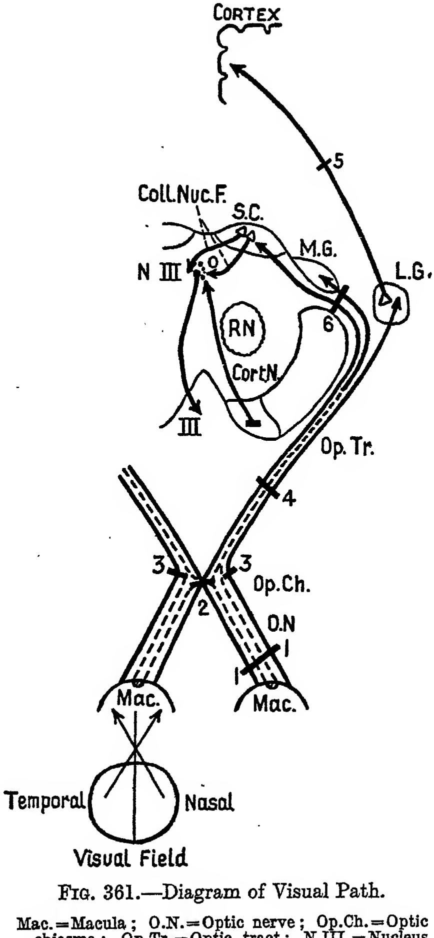


FIG. 361.—Diagram of Visual Path.

Mac.=Macula; O.N.=Optic nerve; Op.Ch.=Optic chiasma; Op.Tr.=Optic tract; N.III.=Nucleus of third nerve; III.=Issuing fibres of third nerve; R.N.=Red nucleus; S.C.=Superior colliculi; Coll.Nuc.F.=Colliculo-nuclear fibres; M.G.=Medial geniculate body; L.G.=Lateral geniculate body; Cort. N.=Cortico-nuclear fibres (from frontal lobes to third nerve nucleus). For significance of numbers 1–6, see text, p. 580.

Central Projection of Visual Fibres.[1]—About one-third of the fibres in the optic tract are derived from the maculæ, and a similar proportion

[1] Le Gros Clark, *Physiol. Rev.*, 1942, *22*, 205.

of the surface of the lateral geniculate bodies is devoted to their reception (Fig. 362). Both anatomical and electrical studies (cf. p. 568) have shown that there is an orderly point-to-point (topographical) projection of the retina, firstly in the lateral geniculate body, and secondly in the occipital cortex.[1]

LATERAL GENICULATE BODIES.—The projection of the retinæ on these bodies is shown in Fig. 362B; the large area allotted to the maculæ is well demonstrated. The grey matter of the lateral geniculate bodies shows six

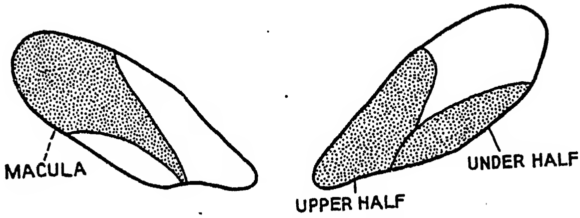


FIG. 362A.—Cross Section of Optic Nerve showing position of Degenerating Fibres.

Left; After lesion of macula. *Right*; After lesion of upper half and lower half of peripheral retina respectively. Degenerating fibres shown dotted.

distinct layers, numbered 1 to 6 (Fig. 363). The fibres from the retina of the contralateral side end in layers 1, 4, and 6; the fibres from equivalent spots on the ipsilateral retina end in the same region but in layers 2, 3, 5. There is thus a fusion of equivalent spots in the lateral geniculate whence fibres pass to the cortex (Fig. 364). Fig. 363 shows that after enucleation of one eye in man, layers 2, 3, and 5, in the lateral geniculate body of the same side show transneuronal degeneration, while layers 1, 4, and 6 remain intact.

CORTICAL REPRESENTATION.—In man the *peripheral* part of the retina

FIG. 362B.—Projection of Retinal Fibres on Lateral Geniculate Body.

Sections through lateral geniculate body showing the relative position of the projections from the macula (white), upper peripheral retina (black), and lower peripheral retina (hatched); sample sections. (After Brouwer and Zeeman, *Brain*, 1926; Fulton, *Physiology of Nervous System*, N.Y.)

is represented well forward on the medial surface of the occipital lobe above and below the calcarine fissure. The *macula* has a very much larger central representation and occupies mainly the posterior part of the medial surface

[1] This can be demonstrated by taking advantage of the occurrence of "transneuronal degeneration" (p. 498). If a spot on the retina is destroyed experimentally in an animal, localized degeneration takes place in the optic nerve and tract and in a perfectly specific region of the lateral geniculate body. The retina can thus be mapped out on the lateral geniculate body. Localized injuries of the visual cortex lead to a *retrograde* degeneration in the nerve cells of the lateral geniculate body; the visual cortex can thus be topographically projected on to the lateral geniculate body. By comparing the results in the two series of experiments the retina may be accurately projected on to the visual cortex. (See Brouwer, *Res. Publ. Assoc. nerv. ment. Dis.*, 1934, *13*, 534.)

(though there is also a forward running tongue). In man the macular representation stops at the occipital pole (Fig. 365).

Each occipital half-vision centre (*e.g.* the left) represents the homolateral (in this case the left) halves of the two retinæ and therefore (as stated) the opposite (right) halves of the field of vision. Furthermore, above the calcarine

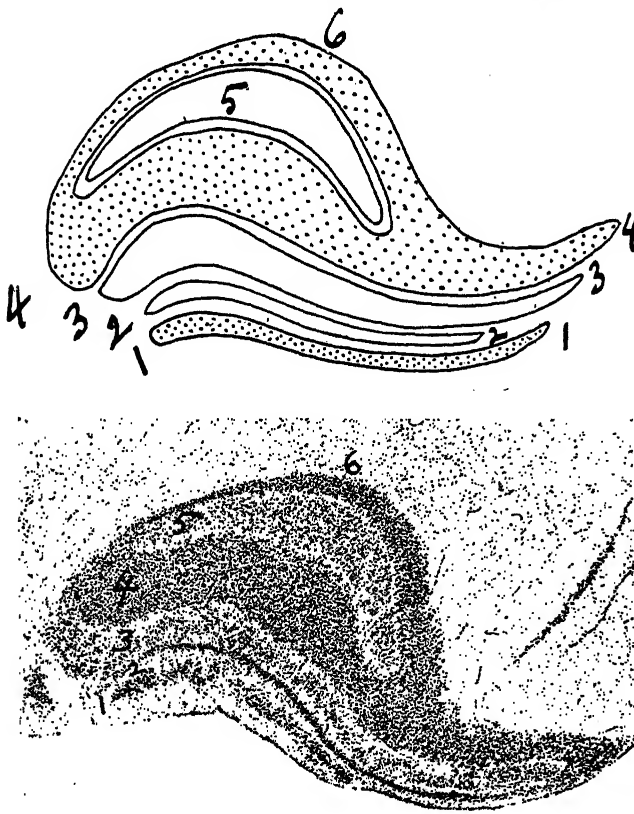


Fig. 363.—Layers of Grey Matter of Lateral Geniculate Body in Man.

Above: Diagram of six layers of grey matter. Crossed retinal fibres end in stippled laminæ (1, 4, 6), and uncrossed fibres in the clear laminæ (2, 3, 5).
Below: Section through grey matter of lateral geniculate after enucleation of the eye on the same side. Note atrophy of cell laminæ 2, 3, and 5. (Le Gros Clark, *Trans. Ophthal. Soc.*, 1942, *62*, 231; *Brit. J. Ophthal.*, 1932, *56*.)

fissure (cuneus), the upper halves of the retinæ (lower halves of the field of vision) are represented; below the calcarine fissure (lingual gyrus), the lower halves of the retinæ (upper halves of the field of vision) are represented.[1]

Both the region of the occipital lobe which surrounds the visual receptive area (*e.g.* area 18) and the posterior parietal region are believed to function

[1] Holmes and Lister, *Brain*, 1916, *39*, 34

as a visuopsychic area. Activity in this region enables the nature of objects seen to be recognized, *e.g.* a pencil, paper, or ball is recognized as such.[1]

Reference is made on p. 669 to the fact that fibres from the frontal eyefield (area 8) pass back to area 18, and that the latter connects with the temporal cortex. It should also be emphasized that area 19 (anterior to area 18 on the lateral surface) is one of the suppressor bands (p. 620). Responses similar to those resulting from stimulating the frontal eyefields can be obtained on stimulating the *occipital eyefields* (in areas 17 and 18) (p. 637).

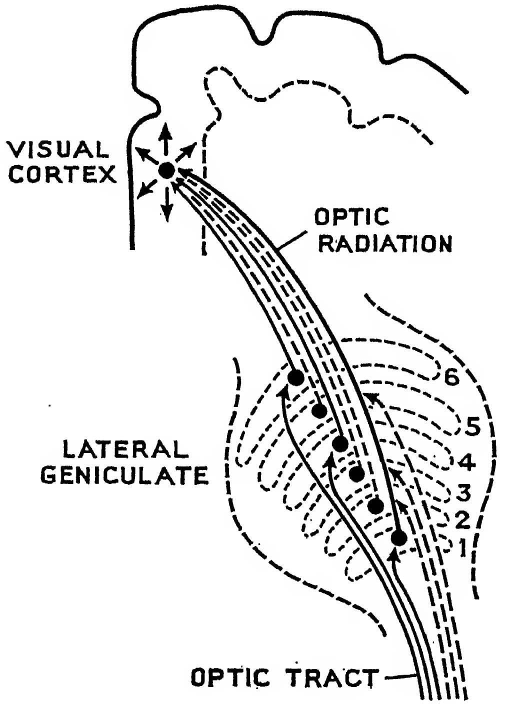


FIG. 364.—Mode of Termination of Fibres of Optic Tract in Six Layers of Lateral Geniculate Body and further Course to Visual Cortex. (After Le Gros Clark, *J. Anat.*, 1941, *75*, 232.)

Continuous lines = crossed fibres; interrupted lines = uncrossed fibres.

The Light Reflex.—The nervous arc employed by the light reflex is probably as follows: The afferent fibres enter the superior colliculi or the adjacent pretectal region from the optic tracts. Here a new relay arises—the *colliculo-nuclear* fibres, which cross both in front and behind the aqueductus Sylvii to reach the most anterior part of the third nerve nucleus. These fibres are probably mingled with the tectospinal tract, which is destined to end in the spinal cord. As the fibres from each retina reach both optic tracts and both superior colliculi, it is found that shining light on one eye causes constriction of both pupils

[1] Le Gros Clark has suggested that the layered arrangement of the lateral geniculate is related to colour vision. The Helmholtz theory of colour vision enunciated in terms of modern sense-physiology supposes that each of the three primary colours, red, green, and violet, stimulates a specific receptor in the retina (presumably a specific type of cone). Each cone connects up with a ganglion cell and thence with an optic nerve fibre. Clark suggests that all the cones from one eye concerned with one colour end in one of the layers of the lateral geniculate; thus the cones for red might end in layer 2 from the ipsilateral eye; those from the contralateral eye would end in layer 1. The new relay from layers 1 and 2 would end in one cortical point (Fig. 355). Activity of what may be crudely called the red, green, or violet pathways *alone* would arouse a sensation of that colour alone. Activity of the three sets of pathways *combined suitably* would arouse all other colours. The sensation of white light would depend on the simultaneous and equal activation of all three colour pathways.

(*consensual* light reflex). The fibres of the third nerve relay in the ciliary ganglion and pass in the short ciliary nerves to the eye.

Convergence-Accommodation Reaction.—During accommodation for near vision the ciliary muscle (on both sides) contracts, the suspensory ligament of the lens is relaxed, and the anterior surface of the lens becomes more convex. At the same time, the eyes converge owing to contraction of both medial rectus muscles, and the pupils are constricted. Convergence-accommodation is to some extent a willed movement, as the object must be definitely looked at before it occurs. It is suggested that visual impulses pass to the occipital cortex and are relayed to the frontal lobes. Fibres arise there which descend in the anterior limb of the internal capsule to reach the medial part of the pes pedunculi. These *corticonuclear* fibres then turn abruptly dorsally through the medial lemniscus to the opposite side, to end in the third nerve nucleus which supplies all three muscles mentioned (see Fig. 361).

Stimulation of area 19 in the occipital cortex leads to constriction of the pupil, the fibres concerned running first to the pretectal area; there is thus another pathway from the cerebral cortex that may be employed in pupilloconstrictor reactions.

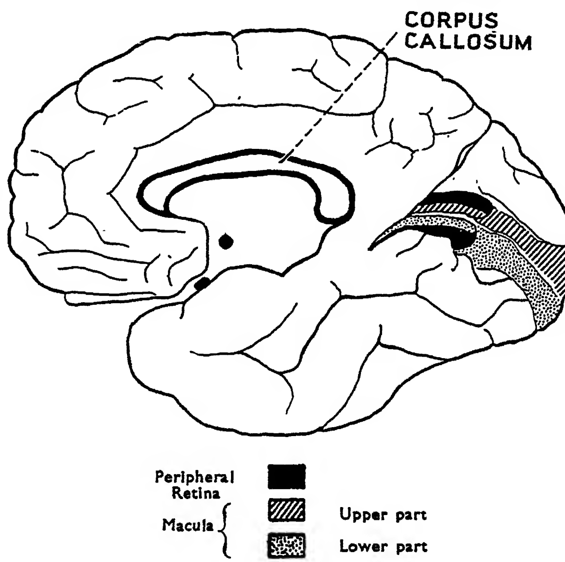


FIG. 365.—Localization of Visual Centres in Occipital Lobe.

Medial surface of the brain in man. The representation of the macula is far greater than that of the peripheral part of the retina. (After Brouwer and Van Heuven. *Res. Publ. Assoc. Res. nerv. ment. Dis.*, 1934, *13*.)

Argyll Robertson Pupil.[1]—This term is applied to a condition in which the pupillary constriction in response to a light stimulus is absent or notably diminished, while the sphincter pupillæ still contracts during convergence-accommodation. From the description given above, it is clear that the only part of the reflex pathway for pupillary constriction which is "private" to the light reflex and is not shared by the convergence-accommodation reaction is the part of the optic tract which enters the superior colliculi, the superior colliculi themselves, and the colliculonuclear fibres (the afferent fibres subserving accommodation pass up to the cortex); the oculomotor nerve is the final common path for both reactions. The Argyll Robertson

[1] Wilson, *J. Neurol. Psychopathol.*, 1921, 2 1

pupil is often associated clinically with lesions in the vicinity of the aqueductus Sylvii and the superior colliculi which would interrupt the "private" light reflex path. As syphilis of the nervous system commonly affects this region, this sign is very frequently found in this disorder.

Effects of Injury to Visual Pathway.—The following definitions may help us to understand the effects of the commoner lesions of the visual tracts. [The numbers in brackets refer to Fig. 361, p. 575.]

Hemianopia: blindness of half the visual field from causes other than retinal. This may be: (i) Bitemporal or binasal, *i.e.* loss of both temporal or both nasal fields of vision. (ii) Homonymous; loss of the right or the left halves of both fields of vision. (iii) Quadrantic: blindness of one quadrant only.

A lesion of the central part of the optic chiasma (2) (*e.g.* from pituitary tumours) where the fibres from the nasal side of both retinæ cross, causes

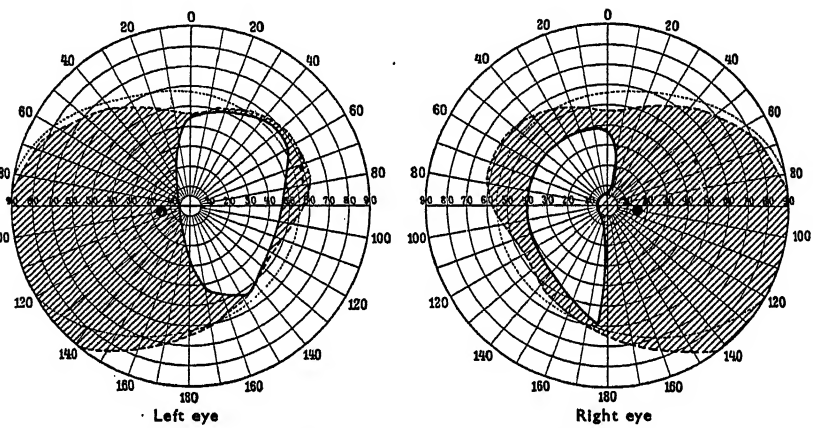


FIG. 366.—Bitemporal Hemianopia.

Loss of both temporal halves of the visual fields in patient with pituitary tumour. The outer dotted line indicates the normal extent of the field of vision; the inner thick continuous line shows the field of vision in the patient. The shaded parts of the chart indicate the degree of visual loss. (H. Zondek, *Krankheiten der Endokrinen Drusen.*)

bitemporal hemianopia (Fig. 366). A lesion of the outer margins of the chiasma (3, 3) may damage the fibres from the temporal sides of the retinæ, and cause binasal hemianopia. Any lesion of the optic nerve, chiasma, or tracts up to the point where the fibres for the superior colliculi leave (1–4), produces loss of the light reflex from the blind side of the retina. A lesion of the lateral geniculate body, optic radiation (5), or occipital cortex produces loss of sight, but the light reflex from the blind side of the retina is retained. A lesion of the optic tracts (4), or their continuation to the cortex (5), causes homonymous hemianopia: *i.e.* a lesion of the left tract or the left visual centre causes loss of the right halves of the field of vision in both eyes.

Incomplete lesions of the visual cortex lead to loss of colour vision; white objects are seen indistinctly, or sensation may be excited only by the more potent type of stimuli such as abruptly moving objects. Lesions of the lateral surface of the brain (areas 18 or 19, *i.e.* in the posterior parietal region (Fig. 430)), leave visual sensibility intact but cause disturbance of

higher visual functions, such as loss of visual orientation and localization in space, impaired perception of depth and distance, loss of visual attention and inability to recognize visually the nature of common everyday objects.

Taste.—Taste receptors (taste buds) are found mainly on the dorsal surface of the tongue but also on the palate, pharynx and tonsils. Four kinds of primary tastes are recognized: sweet, salt, sour and bitter. What we call a taste is a combination of these primary tastes with various smells.

COURSE OF TASTE FIBRES.—(i) *Posterior Third of Tongue.*—The afferent fibres travel in the glossopharyngeal nerve to the medulla.

(ii) *Anterior Two-thirds of Tongue.*—The afferents first travel in the lingual branch of V, then leavei n the chorda tympani; the taste fibres may then take one of two routes: (*a*) continue in the chorda tympani to join the facial nerve and pass centrally in it to the geniculate ganglion (where the cell bodies are found), or (*b*) leave the chorda tympani to pass through the otic ganglion to join the great superficial petrosal nerve and pass in it to the geniculate ganglion.[1]

From the geniculate ganglion the central axons pass in the *nervus intermedius* (afferent and autonomic division of the facial nerve) to end in the pons in the dorsal nuclei of IX and VII; some descend in the related descending tract (*tractus solitarius*) to end in the adjacent nucleus of the tractus solitarius. From this column of grey matter a new relay arises, the fibres crossing to join the opposite medial lemniscus and end in the *thalamus* close to the termination there of the afferent fibres of the *face*. The third relay arises here and ends in the *inferior part of the postcentral cortex* in, or close to, the cortical area receiving the afferents from the face. Changes in the electrocorticogram occur in this area when for example, quinine solutions are placed on the tongue.

It should be emphasized that "common sensibility" on the tongue, (touch, hot, cold, pain) is mediated by receptors the fibres of which pass in the Vth nerve (cell bodies in the semilunar ganglion) to end in the principal sensory nucleus of V and in its long descending nucleus. From what has been said it is clear that all the afferents from the tongue, both of taste and common sensibility, pursue a closely associated course to the thalamus and cerebral cortex. In man, localized lesions of the inferior postcentral region on one side may produce unilateral loss both of taste and of common sensibility on the opposite half of the tongue. Hallucinations of taste have been produced in man by stimulation of the inferior postcentral area.

The older view that the taste fibres end in association with those of smell in the hippocampal region of the cerebral cortex is incorrect.

The sense of taste is not merely a luxury bestowed on us by Providence to enable us to enjoy a wider range of sensual experiences; it also has a limited protective value, causing some (but by no means all) harmful foods to

[1] According to Cushing, removal of the semilunar ganglion of the Vth nerve never causes any loss of taste over the posterior third of the tongue, but after the operation there is a transient loss of taste over the anterior two-thirds. This transient loss of taste is attributed to the post-operative degeneration and *swelling* of the fibres of the lingual nerve (concerned with "common sensibility" in the tongue) which have their cell bodies in the semilunar ganglion. The taste fibres which are carried as "passengers" in the lingual nerve are temporarily pressed on and their conductivity is impaired for a short time. As the lingual nerve fibres atrophy, the fibres of taste recover and regain their function.

is transected, *e.g.* in the midthoracic region (Fig. 367, A) there is an initial period of *spinal shock*; during this time, the distal part of the cord carries out no reflex activities. Muscle tone is absent in the affected region and the muscles are utterly flaccid (cf. p. 690).

After a time the isolated part of the cord resumes functioning as a reflex centre. Some degree of contraction returns slowly, and mainly in the *flexor* muscles of the lower limbs; the extensors, however, remain flabby. In consequence of this unequal distribution of activity between the flexors and extensors, the lower limbs become placed in an attitude of flexion at the knee and physiological flexion (dorsi-flexion) at the ankle. The degree of contraction of the muscles is, as a rule, insufficient to support the weight of the body against the influence of gravity (p. 691). In some patients a greater degree of extensor activity develops as described on p. 693.

It can be readily proved that this flexor position is reflexly maintained; if the appropriate dorsal nerve roots are severed (*e.g.* at 1, Fig. 367), the muscles lose their tone completely.

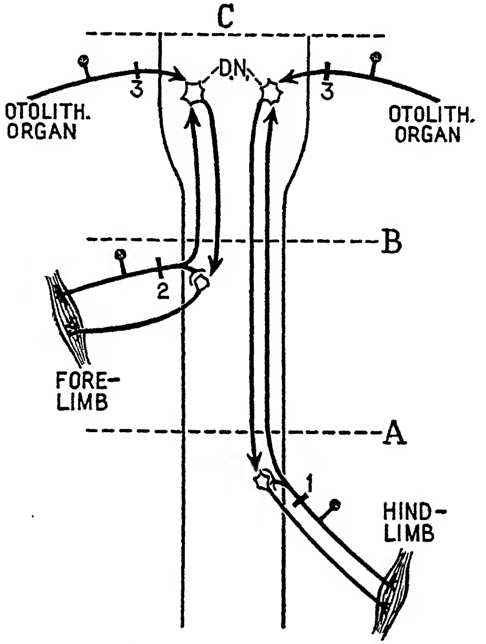


FIG. 367.—Levels of Trans-section to Produce Spinal Preparation and Decerebrate Rigidity (Cat).

A trans-section through the midthoracic (A) and midcervical (B) cord respectively produces a spinal preparation.

A section above Deiters' nucleus (D.N.), in the midcollicular region (C) produces decerebrate rigidity in the cat.

In the case of the hind limb, section of dorsal nerve roots (at 1) abolishes muscle tone; in the fore limb (of the decerebrate cat) the vestibular nerve must be cut also (section at 2 and 3) to abolish muscle tone.

Decerebrate Preparation.—By decerebration is meant carrying out a complete trans-section through the brain stem. As the classical studies on the decerebrate preparation were carried out on the cat, the description that follows is based primarily on findings in this species. But it must be emphasized that as one ascends the animal scale through monkeys and primates, to man, there is an increase in the size of the cerebral hemispheres and in the degree of dominance which they exert over the lower levels of the nervous system; there are correspondingly considerable species differences in the effects of decerebration. In the cat, if the brain stem is divided in the midbrain between the superior and inferior colliculi (*mid-collicular transection*) (Fig. 367, C) decerebrate rigidity develops. With the animal lying on its side the head is drawn back, the four limbs are extended at knee and ankle,[1] the lower

[1] "Physiological extension" of the ankle is what the anatomists call "plantar flexion."

jaw is held up, the back is arched and the tail is lifted up. It is very difficult to alter passively the position assumed by the different parts of the body, hence the name rigidity. Examination shows that many muscles are firmly contracted; such contraction (without movement, under so-called resting conditions) has always been described as *muscle tone* by clinicians. The decerebrate preparation thus displays muscle tone in a high degree, *hypertonus* in fact; but while the preparation is lying down on its side this muscular activity seems devoid of "purpose" or "meaning." But if the decerebrate animal is raised on its legs, it is found that with a little help the preparation stands, too stiffly and imperfectly, it is true (and it tends to topple over), but the distribution of muscular activity now does constitute a purposeful pattern. The muscles which are contracted are now seen to be the antigravity muscles, *i.e.* those which normally maintain the erect position of the body and oppose the action of gravity (*e.g.* limb and head extensors, jaw raisers). In the decerebrate preparation, the muscular hypertonus thus maintains a position or posture, *i.e.* that of "caricatured or exaggerated standing"; the meaning or purpose of tone is thus revealed as a *muscular contraction which maintains a position.* The position at any moment will vary with the *degree and distribution of the tone.* A detailed study of decerebrate rigidity gives a good deal of insight into the mechanisms responsible for muscle tone.

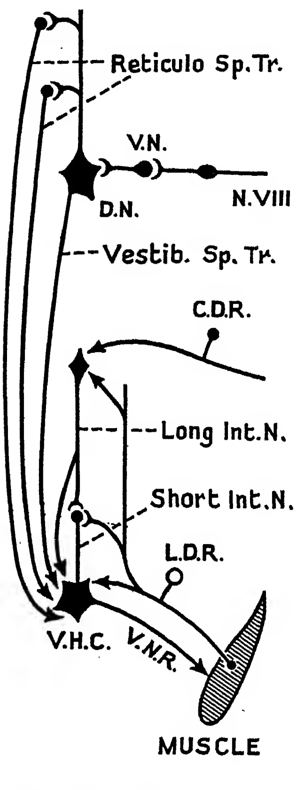


FIG. 368.—Reflex Arcs responsible for Decerebrate Rigidity (Cat).

L.D.R., lumbar dorsal nerve root. V.N.R., ventral nerve root. C.D.R., cervical dorsal nerve root. V.H.C., ventral horn cell. Short Int. N., short interneurones. Long Int. N., long interneurones. Vestib. Sp. Tr., vestibulospinal tract. D.N., Deiters' nucleus. N.VIII, Vestibular division of VIIIth nerve. V.N., Vestibular nucleus. Reticulo Sp. Tr., reticulospinal tract.

Reflex Arc of Decerebrate Rigidity.—(1) The muscle tone producing the rigidity is reflex in origin; in the hind limb, for example, the rigidity disappears on section of all the dorsal nerve roots. The afferent impulses concerned do not come from the skin but from the muscles and tendons: (i) removal of the skin over a limb does not affect the rigidity; (ii) if all the nerves to the limb are cut except the nerve to quadriceps, tone in that muscle persists; this nerve contains, of course, the afferent fibres from the muscle as well as the efferent fibres to the muscle. (iii) If the dorsal nerve roots which receive the afferents from the quadriceps are cut, tone in the muscle disappears. Tone in a muscle thus depends on afferent impulses from itself, *i.e. tone is a proprioceptive reflex.*

(2) The afferent impulses concerned are set up by the slight *stretch* to which the various antigravity muscles are subjected; if the stretching force

is removed by cutting off the quadriceps tendon from its attachment to the bone, tone disappears. Muscle tone is thus due to a *stretch reflex* (p. 588). Impulses set up by stretch stimulate the ventral horn cells to discharge to the muscles concerned.

(3) Decerebrate rigidity is a *release phenomenon*; the mid-collicular transection removes the cerebral hemispheres leaving only part of the brain stem and the spinal cord in control of the musculature. It cuts off the descending tracts from the cerebral cortex (both pyramidal and extra-pyramidal tracts), from the diencephalon (globus pallidus, thalamus, hypothalamus) and from the red nucleus and reticular nuclei; as is explained on p. 593 some of these tracts normally *increase*, others *decrease* muscle tone by their influence on the lower motor neurones. The "rump" of lower nervous system which is released by decerebration from the modifying influence of the higher levels produces reflexly the abnormal pattern of tone which constitutes decerebrate rigidity.

(4) If the trans-section is carried out further back, the rigidity persists until the upper medulla is reached; when the *vestibular nuclei are destroyed*, the rigidity disappears and is replaced by tonelessness (flaccidity).

(5) The complete reflex arc, *e.g.* for hind limb tone, is as follows (Fig. 368): Afferent fibres from the hind limb muscles enter the lumbo-sacral dorsal nerve roots and pass directly to the motor neurones; collaterals also pass to short internuncials in the ventral horn. Ascending dorsal root fibres give collaterals to internuncial cells in higher segments of the cord which give rise to descending long propriospinal paths (that descend to the lumbo-sacral region). Still longer ascending fibres reach the medulla and pons and connect with the vestibular nuclei and the nuclei of the reticular formation; from these cells, descending tracts arise, *i.e. vestibulospinal* from the vestibular nuclei, *reticulospinal* from the reticular grey matter. The long propriospinal and the tracts from the brain stem descend to end round ventral horn cells and the adjacent internuncials. After a midthoracic spinal transection, the stretch reflex set up in the hind limb muscles operates solely through spinal centres and produces a weak contraction, mainly of flexor muscles; in the decerebrate preparation the basic spinal reflex is reinforced and modified by impulses that pass through the upper cord and brain stem and descend in the long propriospinal, vestibulospinal and reticulospinal paths. This complex reflex arc, released from the influence of the higher levels, sets up decerebrate rigidity.

(6) The decerebrate posture in *man* is shown in Fig. 372; a similar posture is seen in primates.

Motor Neurone Discharge in Decerebrate Rigidity.—As explained on p. 502 each motor nerve impulse produces a single muscle action potential. The rate of discharge of an individual motor neurone can, therefore, be determined simply by inserting a fine needle electrode into the related motor unit (*i.e.* the clump of muscle fibres supplied by the motor neurone) and recording the muscle action potentials. The motor neurone discharge rate in decerebrate rigidity is found to be low, *i.e.* 5–20 per sec. Now mammalian muscle stimulated at this rate passes into a state of partial tetanus (Fig. 300); therefore if all the motor neurones were discharging synchronously at this rate, the muscles would be in a state of tremulous contraction and not maintaining a uniform tension as is actually the case. It is concluded therefore

that in decerebrate rigidity the motor neurones are discharging *asynchronously*; at any instant the motor units are contracting *out of phase* and the algebraic sum of their activity is a steady pull (Fig. 301). The degree of tone in any muscle depends on (i) the number of motor neurones in action and (ii) their discharge frequency. When fewer motor neurones discharge and at a lower frequency the resulting muscle tension decreases (hypotonus); when more motor neurones discharge and at a higher frequency, increased muscle tone (hypertonus, spasticity, rigidity) results. Still greater motor neurone discharge, further enhancing muscle tension, produces a movement.

When skeletal muscle is maintaining tone, it does *not fatigue* and has a relatively *low metabolism*. The explanation is as follows: The tension exerted by tonic muscle is far smaller than that developed in the same muscle when it is stimulated at a high rate (*e.g.* 100 times per second) through its motor nerve to produce a *simultaneous and complete tetanus of all the muscle fibres*. Thus the quadriceps of the cat develops 6 kg. of tension when maintaining tone and 30 kg. when stimulated tetanically. As the work of tonic muscle is relatively small, its oxygen consumption (metabolism) is similarly low. Its unfatigability is due to the fact that it is possible to increase the blood flow to the tonic muscle so as to supply it with its full nutritional requirements. In the case of fully tetanized muscle, fatigue rapidly develops as it is not possible for the circulation to cope with the great demands of the muscle.

LENGTHENING AND SHORTENING REACTIONS.—(i) If an attempt is made to bend the knee of a decerebrate animal, considerable resistance is encountered. The attempt increases the stretch of the muscle and so elicits a further reflex contraction. But if one persists, the excessive stretch stimulates other, unidentified, receptors which reflexly inhibit the motor neurones; their discharge then ceases, the muscle fibres relax, and the limb can be placed in any desired position of flexion. This is the *lengthening reaction*, so called because the muscle fibres have been reflexly lengthened.[1]

(ii) If, starting from the new flexed position, the knee is extended again, it tends to stay in the new position; this is the *shortening reaction*. The mechanism involved is probably as follows: When the knee is passively extended and the limb support is withdrawn, the leg tends to droop by reason of its own weight; as a result a stretch reflex is set up and the additional tension of the muscle enables it to maintain the new posture.

The shortening reaction is grafted on to many reflexes. Thus, if the crossed extensor reflex (p. 537) is produced in a *de-afferented* quadriceps recording isotonically, on cessation of the stimulus, the tension is maintained for a brief period (after-discharge), and then rapid relaxation occurs (Fig. 369, d). If the afferents from the muscle are intact, when the stimulus ceases there is an initial slight decline from the maximum level; the tendency for the limb to fall sets up reflexly a shortening reaction which results in the limb being maintained in its new position for a prolonged period (Fig. 369, a).

TYPES OF MUSCLE FIBRE EMPLOYED.—In vertebrates no skeletal muscles are used exclusively in postures: the same muscle is used at different times for movement or for posture. In any complex movement some groups

[1] Sherrington has demonstrated the presence of inhibitory afferents in the nerves to the quadriceps; if one of the four branches which supply the muscle is cut and the central end stimulated, the rest of the muscle is inhibited.

of muscles are executing the movement, while others are chiefly maintaining a posture. Thus, in the scratch reflex of the dog, three of the limbs and the head and neck are kept in a characteristic posture, while the fourth limb carries out the scratching.

There are two main kinds of fibres in skeletal muscle: (i) *red* fibres, rich in sarcoplasm, with poorly marked transverse striations, and nuclei scattered through the substance of the fibre; the colour is due to the presence of *myohæmoglobin* (cf. p. 204); (ii) *pale* fibres, with well-marked cross striations, scanty sarcoplasm, and nuclei limited to the under-surface of the sarcolemma. In addition, numerous intermediate types of fibre are found. They all have a similar motor innervation. In man, most muscles

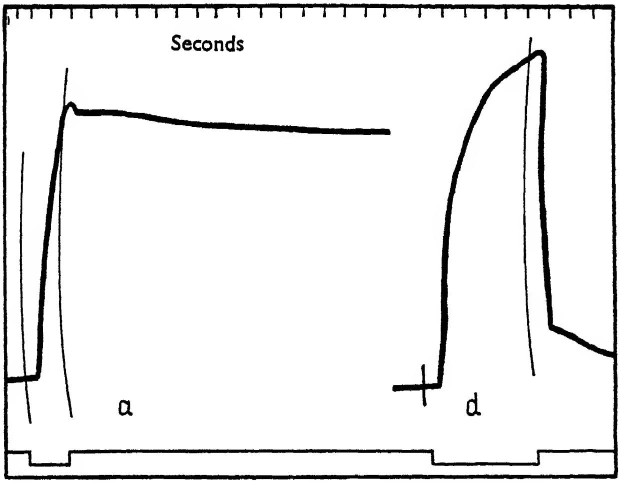


FIG. 369.—Shortening Reaction grafted on to Crossed Extensor Reflex.

Records from above downwards are: time in seconds; isotonic contraction of vastocrureus; signal line showing by its descent period of stimulation of afferent nerve in opposite limb to elicit crossed extensor reflex. *a*, Shows response of muscle with its own afferent roots intact; *d*, Same muscle with afferents cut 2 hours previously. Note (in *a*) that following the after-discharge "plateau" there is a slight decline and then there is appended a prolonged shortening reaction; this is absent in *d* proving that the shortening reaction depends on proprioceptive impulses from the responding muscle itself. (Sherrington, *Quart. J. exp. Physiol.*, 1909.)

contain a mixture of both red and pale fibres, though the red predominate in some (*e.g.* soleus), and the pale in others (*e.g.* gastrocnemius). Probably the red fibres are better adapted for maintaining postures and the pale for executing rapid movements. The red fibres are more primitive in structure and function; they contract slowly after a longer latency, the duration of the contraction being three times that of the more quickly acting and more highly differentiated pale fibres. Fibres of similar speed of contraction generally form a group which is sharply delimited from other groups; the deeper muscle "heads" are usually composed of the more slowly contracting red fibres. These latter have a *lower threshold to stretch* than the pale fibres. They probably come into action *first* or alone when mild stretch is employed; but with greater stretch the pale fibres come into operation too. On the other hand, the faster acting pale fibres would be used first during movements.

But it must be emphasized that both kinds of fibre may be employed under suitable circumstances in either movement or posture.

Stretch Reflexes.[1]—These reflexes, which are of fundamental importance in the production of muscle tone, must now receive fuller attention. The method of study is as follows: In the decerebrate animal the quadriceps muscle is isolated from its attachment and fixed, by means of a steel hook in the patellar tendon, to a stiff spring, to which a light lever is attached; its movement is recorded photographically on a moving plate. The record obtained is an *isometric* one, *i.e.* when the muscle contracts it is only permitted to shorten to a *minimal degree*; under these conditions the *increase in tension* (not the shortening) of the fibres is recorded. The hind limbs and pelvis are firmly fixed to a rigid table, which can be lowered to any degree and at any rate; the muscle can thus be subjected to carefully controlled stretch.

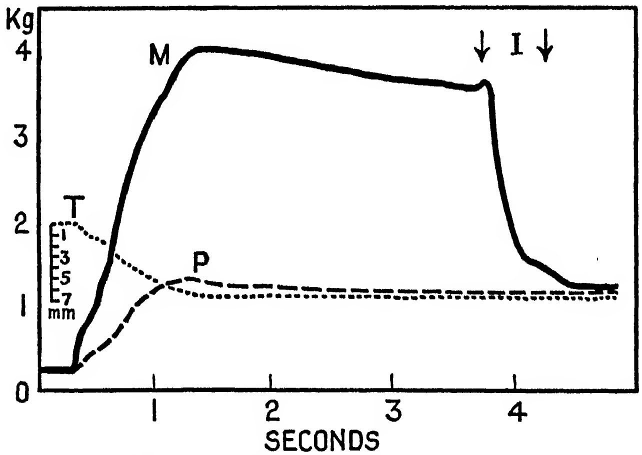


Fig. 370.—Stretch Reflexes. Response of Intact and Paralysed Muscle to Stretch. (Liddell and Sherrington, *Proc. roy. Soc. B.*, 1924.)

Ordinates: Muscle tension in kg., and stretch of muscle in mm.
Response of Knee Extensor Muscle.—T = Stretch of 6·5 mm. applied; M = Innervated Muscle; P = Paralysed muscle; I = Inhibitory afferent nerve stimulated, which abolishes the stretch reflex, and tension in M falls to level of P. Myograph multiplies tendon movement 62 times.

Response of Denervated Muscle.—If a *paralysed* muscle (*i.e.* with the motor nerve cut) is stretched, a very small increase in tension results; the response obtained resembles that given by stretched indiarubber (Fig. 370, P).

Response of Innervated Muscle.—If the nerve supply of the muscle is intact, stretch gives rise to a considerable increase in tension (Fig. 370, M).

(i) The response is *reflex* in character as it is abolished by section of the dorsal nerve roots which carry the afferents from the muscle. A stretch of several millimetres (*i.e.* less than 1 per cent. of the initial length of the muscle) may produce a reflex contraction of 2 kg. tension. The reaction though most marked in the decerebrate preparation, is also obtainable in the thalamic and the intact animal.

(ii) When the stretch is applied, rapid development of tension occurs.

[1] Liddell and Sherrington, *Proc. roy. Soc. B.*, 1925, *96*, 212; *97*, 267.

While the stretch is maintained, a fairly steady tension is kept up, *e.g.* for 30 minutes, without fatigue. When the stretch ceases, the muscle relaxes. Muscle thus reflexly responds to stretch by an active contraction, which antagonizes the stretching force, produces equilibrium, and so prevents further elongation of the muscle.

(iii) The stretch reflex can still be obtained after cutting away the tendon and attaching the muscle directly to the myograph; obviously, then, the reflex can be set up by muscle afferents; it is probable however that normally, afferents from tendon are also concerned.

(iv) If the quadriceps is divided into four segments, only the part which has been stretched contracts; the response is thus intensely *local* in character.

(v) Typical stretch reflexes are a property of the *antigravity* muscles only; stretch of the flexor muscles gives rise to a short-lived *twitch-like* response, but not to the characteristic maintained contraction.

(vi) The stretch reflexes may be readily inhibited, *e.g.* by stimulating any afferent nerve in the same limb, or by pulling on the antagonistic flexor muscles, even while the stretching force is being applied.

(vii) The stretch reflexes vary with the *initial state of tone* in the muscle under observation: thus, if the quadriceps is stretched while it is under reflex inhibition, it gives the response characteristic of paralysed muscle. If the quadriceps muscle is reflexly caused to contract (by eliciting the crossed extensor reflex), and is then subjected to stretch, it gives a larger response than a resting muscle (Fig. 371). The explanation is that a given degree of stretch elicits a greater discharge of impulses from the receptors in a contracted than in a relaxed muscle.

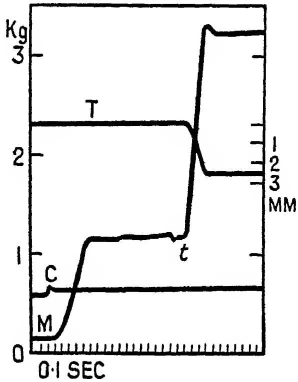


FIG. 371.—Stretch Reflex of Contracting Muscle. (Liddell and Sherrington, *Proc. roy. Soc. B.*, 1924.)

Ordinates: Muscle tension in kg. (left-hand scale); muscle stretch in mm. (right-hand scale).
Records from above downwards. T=stretch of muscle; C=signal; M=muscle tension.
At C begin stimulation of afferent nerve on opposite side evoking reflex contraction (crossed extensor reflex). At *t*, 2·5 mm. stretch (right-hand scale) applied (shown on line T) evoking marked stretch reflex from the contracting muscle.
Time in 0·1 second. Tendon movement multiplied 62 times.

(viii) In the *spinal* animal, the stretch reflex is considerably modified; stretch of the quadriceps only produces a transient and not a sustained contraction. The response resembles the familiar *knee-jerk*, which is simply a fractionated or abbreviated stretch reflex.

(ix) In the *chronic* spinal animal (for man cf. p. 693) *sustained* stretch reflexes can also be elicited. This last observation proves that the stretch reflex depends essentially on a spinal reflex arc. But the fact that stretch reflexes are so much more readily and strikingly elicited in the decerebrate animal demonstrates that these reflexes (as explained on p. 585) are reinforced (*facilitated*) by impulses coming down in the long propriospinal, vestibulospinal and reticulospinal paths.

POSITIVE SUPPORTING REACTION.—Decerebrate rigidity basically depends on a *harmoniously operating group of stretch reflexes*. These highly *localized*

reflexes produce contraction of the antigravity muscles and *reciprocal inhibition of the antagonistic muscles.* The resulting posture is reinforced and modified by the positive supporting reaction; this is a remarkable irradiating reflex which produces *simultaneous contraction* of extensors and flexors of a limb converting it into a solid rigid pillar, well adapted to maintaining the body weight. The positive supporting reaction, for some unknown reason, is most easily elicited after complete removal of the cerebellum. The procedure is as follows: press against the pads of the fingers or toes and dorsiflex the hand or the foot; the afferent impulses arise from the stimulated skin and from the muscles (mainly the interossei) which are stretched. *All the muscles of the limb reflexly contract,*[1] *i.e.* both the protagonists and the antagonists. In this instance there is reciprocal *contraction* of the antagonists (cf. p. 649).

The positive supporting reaction is readily abolished as follows: passively flex the digits and hand in the forelimb, or plantar flex the toes in the hind limb. As a result, *all* the muscles in the limb are reflexly inhibited and the part becomes loosened and free to be employed in the execution of movements.

The decerebrate "stance" is thus maintained as follows: When the decerebrate animal is put on its legs, the knees tend to bend under the weight of the body; the consequent stretch of the quadriceps reflexly sets up its contraction, which is maintained so long as gravity is acting; and similarly for the posterior neck muscles and the other antigravity muscles. Positive supporting reactions are set up from the feet in the way already described.

SEGMENTAL STATIC REACTIONS.—The decerebrate animal also possesses reflex mechanisms for adjusting the position of one limb in relation to alterations in the state of another. The *crossed extensor reflex* (p. 537) is a case in point; impulses from one leg reflexly produce extension of the opposite limb. Another example is the *shifting reaction*: flex (say) the right limb and allow the body to veer to the right; owing to stretch of the adductors of the left limb, the right limb is reflexly caused to extend.[2]

Attitudinal Reflexes in the Decerebrate Preparation.[3]—The above discussion shows that the postural activities of the decerebrate preparation represent a great advance on those found in the spinal preparation. The posture is now a co-ordinated one of the whole body instead of being limited to the hind limbs (or to all four limbs in a high spinal transection), and the degree of tone is adequate to maintain the upright position. We shall now see that in the decerebrate animal the posture of the trunk and limbs can be adjusted: (i) in accordance with alterations in the *position of the head in space,* and (ii) by *changing the position of the head relative to the trunk.* In the first case the afferent impulses arise solely from the otolith organ of the vestibule (tonic labyrinthine reflexes [4]); in the second case additional afferent impulses come from the neck muscles (tonic neck reflexes [4]). The new position reflexly imposed on the body persists for as long as the new position of the head is maintained.

[1] The *extensor thrust* reflex (p. 694) may represent a "fragment" of this more genera reaction.

[2] Stretch of the quadriceps also causes contraction of the opposite quadriceps (Phillipson's reflex): *e.g.* forced flexion of one leg at the knee causes extension of the opposite leg and foot (cf. p. 694).

[3] Magnus, *Körperstellung,* Berlin, 1924.

[4] *i.e.* reflexes from labyrinth or neck respectively which modify tone or posture.

(1) Tonic Neck Reflexes.—To study these separately, bilateral extirpation of the labyrinths is carried out. The neck reflexes are set up by *alterations of the position of the head relative to the body.*

(i) Ventroflex the head: the fore limbs flex and the hind limbs become more extended.

(ii) Dorsiflex the head: the fore limbs extend and the hind limbs flex. The purpose of these responses seems obvious: the position of the body is being adapted, *e.g.* " for looking under a shelf, or looking up to a shelf."

(iii) Press ventralwards on the lower part of the cervical vertebral column: all four limbs flex (*vertebra prominens reflex*), as " in an animal crawling into a hole."

(iv) Rotate or incline the head in various directions: to simplify description, the limbs on the side to which the jaw is turned are called " jaw limbs " ; the limbs to which the vertex is turned are called " skull (or vertex) limbs." In general it may be said that the *jaw limbs extend* (to support the weight of the head) and the *skull limbs flex.*

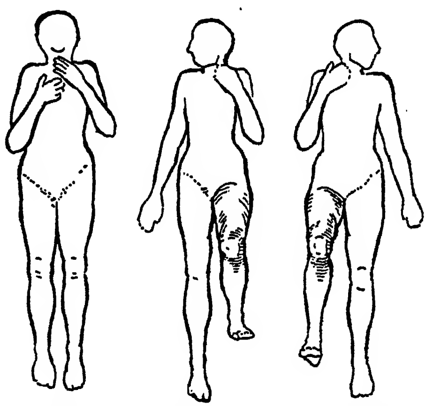


Fig. 372.—Posture of Limbs in Decorticate Man, and the Effect upon them of Rotation of the Head. (After Walshe.)

The centre for reflexes (i), (ii), (iii) is in the upper cervical region of the spinal cord; the afferent impulses pass in the dorsal roots of C 1–3 and come chiefly from the muscles of the back of the neck. The descending paths are the long propriospinal tracts (Fig. 373). The centre for (iii) is in the lower cervical region.

(2) Tonic Labyrinthine Reflexes.—These are studied after section of the dorsal nerve roots of C 1, 2, 3 or after immobilizing the head, neck, and upper thorax by means of a plaster jacket (to prevent neck reflexes from coming into play). The labyrinthine reflexes are due to *alterations in the position of the head relative to the horizontal plane.*

(i) If the animal is placed in the *supine* position, maximum tone is present in the antigravity muscles. (ii) In the *prone* position, with the snout 45° below the horizontal plane, tone in the extensor muscles is reduced to a minimum; in intervening positions intermediate grades of tone are present. The purpose of these reactions is not very clear; they disappear after section of the vestibular nerves. The receptors are in the otolith organ as shown by the following experiment. If an anæsthetized guinea-pig is centrifuged at high speed, the otolithic membranes are detached (as proved by microscopic examination), but the ampullæ of the semicircular canals are unharmed: the labyrinthine reactions are, however, abolished. Alterations in the position of the otoliths (cf. p. 598) thus reflexly modify tone in the muscles of the limbs. The centres for the labyrinthine reactions are the vestibular and reticular nuclei; the descending tracts employed are the vestibulospinal and reticulospinal (cf. Figs. 368, 373).

After unilateral labyrinthine extirpation (p. 601), the unopposed activity of the intact otolith organ results in flexion and rotation of the head to the side of the lesion; this in turn evokes secondary modifications in the position of the limbs through the operation of the neck reflex. As a result the jaw limbs extend and the skull limbs show diminished extensor tone.[1]

The attitudinal reflexes can be easily demonstrated in normal animals and man where they are carried out more smoothly because the flexor muscles also participate in the reactions. A cat sees a piece of meat in the air; the head extends, the fore limbs extend, the hind limbs are unaffected or flex, but can be suddenly extended when the animal wants to spring.

Posture in the Decorticate Preparation.—The whole cerebral cortex is removed but the basal ganglia and the brain stem are left intact. The posture adopted by such a decorticate preparation varies with the species and the conditions under which the animal is examined.

In the dog or cat the posture when the animal is on its feet is a *normal* one (unlike the exaggerated standing posture of decerebrate rigidity); muscle tone is, therefore, normally distributed and is present in appropriate measure in the flexors as well as in the antigravity muscles; *walking movements* can be reflexly performed. If the preparation is suspended in the air, strong hyperextension of the limbs develops. In the decorticate monkey, however, tone is gravely disturbed; there is full extension of the hind limbs and a semi-flexed posture of the fore limbs; walking movements do not occur. The findings in decorticate man are very similar and are illustrated in Fig. 372. The legs are fully extended; the arms lie across the chest, semiflexed at the elbows, the forearms slightly pronated, and the wrists and fingers flexed.

In decorticate man typical neck reflexes can be elicited. When the head is rotated to the right, the right arm extends at the elbow and becomes abducted; the left arm becomes fully flexed with the hand touching the neck; the right leg is extended and the left leg is flexed (Fig. 372). The usual rule is followed here: the skull limbs flex and the jaw limbs extend. The human decorticate posture is of great interest because an *almost identical posture of the arms and legs is found in hemiplegia* (p. 642); it is probably largely due to removal of the influence of the inhibitory pathway arising in the cortical suppressor band 4s (cf. p. 623).

Righting Reflexes in the Decorticate Preparation.—By means of the righting reflexes the decorticate cat or rabbit can bring its head right way up and get the body into the erect position under all circumstances. If the animal is laid on its side or on its back, the head at once rights itself, the body follows suit, and finally the animal resumes the upright posture. The decerebrate animal, though it can remain insecurely in the upright position if put there, can never actively assume that position. The righting reflex consists of a chain of reactions following one another in an orderly sequence as each reaction produces its successor.

[1] When tonic labyrinthine and neck reflexes are simultaneously evoked, they produce the algebraic sum of the separate responses. Thus, if the head is dorsiflexed in a decerebrate animal (with intact labyrinths), labyrinthine impulses produce increased tone in all four limbs, while the proprioceptors from the neck tend to extend the fore limbs and flex the hind limbs. The actual result observed is extension of the fore limbs (as both reflexes tend to increase extensor tone) and little change in the hind limbs (because the two reflexes are exerting antagonistic influences).

(1) Labyrinthine Righting Reflex.—With the animal's head in the lateral position impulses arise in the saccules which lead reflexly to righting of the head.

(2) Body Righting Reflex.—With the animal on its side, the side of the trunk in contact with the bench is undergoing constant stimulation, while the other side in contact with air is not. This *asymmetric* stimulation of the deep structures in the body wall also reflexly rights the head. The head can thus be righted even after double labyrinthectomy.

If a *labyrinthless* animal is laid on its side and a weighted board is placed on the upper side of the animal so that equal streams of impulses pass up from both sides of the trunk, the head falls back into the lateral position.

(3) Neck Righting Reflex.—The reflexes just described act primarily on the *neck* muscles and right the head. The trunk, however, remains as before in the lateral position, so that the neck is twisted. This evokes a further reaction—the *neck righting reflex*—which brings the thorax and lumbar region successively into the upright position. If righting of the head is prevented, impulses from the body surface may cause righting of the body directly.

(4) Limbs.—The appropriate posture of the limbs is largely attained by impulses arising in the limb muscles themselves. The righting reflex can be well demonstrated in the intact cat if it is blindfolded and dropped with the legs pointing upwards. The cat turns itself round with great speed and alights gently on all fours.

The chief centre for this group of righting reactions is in the neighbourhood of the red nucleus (Fig. 373).

(5) Optical Righting Reflexes.—In animals with the *calcarine cortex intact*, righting of the head is also brought about reflexly by means of optical impulses. In the intact cat, dog, or monkey, after denervation of the labyrinths and the neck muscles, righting cannot take place if the animal is blindfolded, but is successfully carried out if the eyes are open. The centre for the optical righting reflex is in the visual (calcarine) cortex, whence impulses pass ultimately to the neck muscles to right the head. In man the optic righting reflexes are far more important than the labyrinthine.

General Regulation of Posture.—Posture is determined by the degree and distribution of muscle tone, *i.e.* of muscular contraction; it depends, therefore, on the *pattern of discharge* of the motor neurones which supply the muscles. Motor neurone activity maintaining tone is reflexly regulated, the basic spinal reflex (stretch reflex) being reinforced and modified by the positive supporting reaction and other reflexes considered on pp. 589–593. When longer reflex arcs are available through the brain stem and basal ganglia, the more complex and efficient postures of the decerebrate and decorticate preparation appear. Afferents which modify the discharge of the centres which give rise to the long propriospinal, vestibulospinal or reticulospinal tracts, reflexly modify posture; hence the adjustments of posture produced by neck reflexes, labyrinthine reflexes and righting reflexes. That the *cerebral cortex*, through its projection fibres, modifies posture is clear from the occurrence of optical righting reflexes. Two groups of fibres arise in cells in the brain stem and descend to end at the spinal motor neurones: some are *facilitatory* (*e.g.* reticulospinal fibres, vestibulospinal tract) and increase muscle tone, the others (*e.g.* other reticulospinal fibres) are *inhibitory*

and decrease muscle tone. The nuclei of origin of these fibres in the brain stem are acted on by *cerebellar* efferents which thus modify tone. The inhibitory pathway is controlled by the suppressor bands in the cerebral cortex ; removal of the suppressor area 4s or decortication results in spasticity and abnormal postures (p. 624). The facilitatory pathway is controlled by cortical area 6

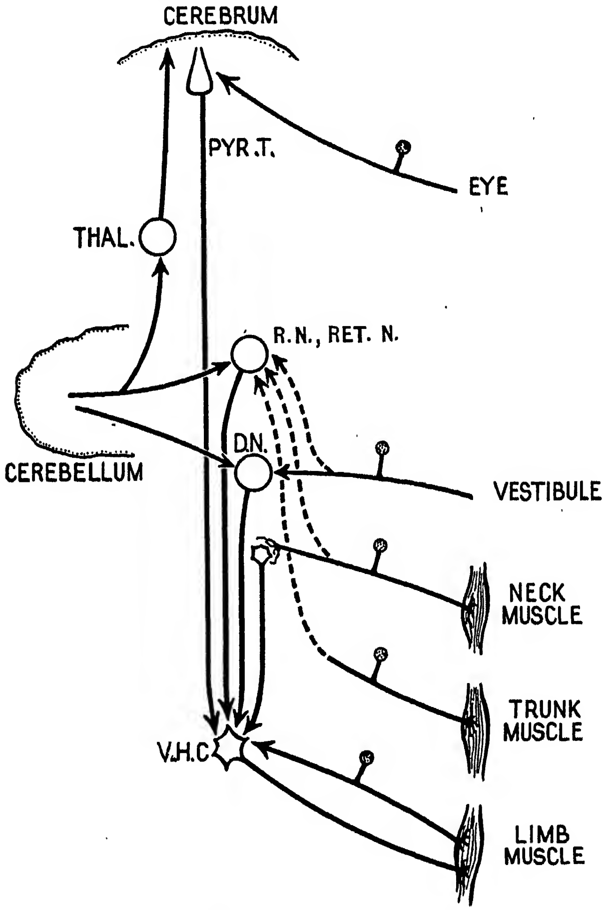


FIG. 373.—Regulation of Posture.

Thal., Thalamus. Pyr.T., Pyramidal tract. R.N., Red nucleus. Ret.N., Nuclei of reticular formation of brain stem. D.N., Deiters' nucleus. V.H.C., Ventral horn cell. (For more details see Figs. 367, 368, 383, 384 400, 408, 424, 425.)

whence fibres pass to relay in the globus pallidus, thalamus and hypothalamus to end in the reticular nuclei (p. 629).

Regulation of normal posture in the intact animal depends on the *integrated activity of all the reflex mechanisms* to which reference has just been made (Fig. 373). Each ventral horn cell is a convergence point ; on it converge fibres from many dorsal nerve roots and descending fibres from all levels of

the brain and spinal cord. Some of these fibres are excitatory, others are inhibitory to the motor neurones. When the activities of all these converging paths are properly co-ordinated, normal posture results ; if any contributing path is destroyed or functions abnormally, posture is deranged, often in a manner so characteristic as to be of great diagnostic value to the clinician. Finally it must be emphasized that posture is the basis of movement ; all movements start and end in a posture, they grow out of a posture [1] and return to it.

Mechanism of Standing in Man.—The comfortable stance of a normal man is quite unlike the imperfect caricature of standing displayed by a decerebrate cat ; it is hardly surprising to find that the detailed mechanisms employed differ though fundamental similarities persist. The activity of the muscles in man can be conveniently studied by : (i) inspection ; (ii) palpation ; (iii) the use of surface electrodes placed on suitable parts of the skin overlying the muscles which are being examined (Fig. 420) ; (iv) concentric needle electrodes inserted into the muscles (p. 648). Using these methods it can be shown that when a man is comfortably balanced in the upright position there is remarkably little activity taking place in the muscles of the trunk and thighs (Fig. 421). The explanation of this surprising finding is that the disposition of the skeleton, the ligaments and the soft parts is such that a *momentary insecure* balance can be maintained *passively*. This passive erect posture in the absence of *all* muscular activity is, as stated, momentary and insecure, and the person would immediately fall if muscular activity did not develop. (A person whose muscles are paralysed cannot stand.) But as soon as the man begins to fall, reflex compensatory muscular reactions set in which restore the state of balance : the muscular contraction then ceases till the next deviation from the erect position occurs. A standing man can fall in any direction : forwards, backwards or sideways. The muscles which oppose the fall are acting as antigravity muscles ; depending on the direction of the fall, any of the muscles of the trunk or legs act as antigravity muscles. Thus when the body sways forwards the extensors of the trunk and the flexors of the leg contract sufficiently to restore the balance ; when the sway is backwards the recti abdominis and the leg extensors contract ; when the sway is sideways the contralateral external oblique muscle responds (Fig. 423). These responses are reflexly produced, partly as a result of impulses from stretch receptors in the trunk and legs and partly from the receptors in the head, mainly the eyes. It is a common experience that with the eyes closed the upright stance is less steady and that more swaying of the trunk occurs. This observation shows that visual afferents are concerned in the reflex maintenance of the upright stance in man.

These questions are considered further on p. 650.

THE VESTIBULAR APPARATUS [2]

The vestibular apparatus (labyrinth) is a complex sense organ which is stimulated by (*a*) gravity and (*b*) rotation movements. It plays an important

[1] Or as the Chinese would say, "a perfect Yin-state passes over into new Yang-activity." (Toynbee, *A Study of History*).

[2] Magnus, *Körperstellung*, Berlin, 1924 ; Croonian Lecture, *Proc. roy. Soc. B.*, 1925, *98*, 339.

rôle in postural activity; it gives rise to afferent impulses which reflexly *adapt the position of the trunk and limbs to that of the head* and enable the *erect position of the head* and the normal attitude of the body to be maintained. Impulses from the vestibule also reach the cerebral cortex and subserve the recognition of the position and movements of the head.

Anatomy.—The vestibular apparatus consists of the three semicircular canals and the otolith organ (the saccule and utricle) (Fig. 374).

(i) The *canals* are the lateral (horizontal), superior, and posterior, each being in a different plane at right angles to the others. The left superior canal is in the same plane as the right posterior canal, and vice versa. The membranous canals contain the endolymph and are enclosed in bony canals,

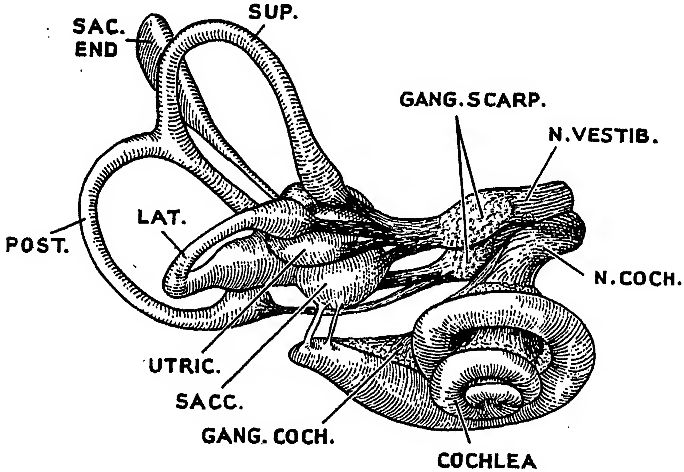


Fig. 374.—Anatomy of Vestibular Apparatus. (Modified from Hardy, *Anat. Rec.*, 1934, *59.*)

N.Vestib. = vestibular nerve.
N.Coch. = cochlear nerve.
Gang. Scarp. = Scarpa's (vestibular) ganglion.
Gang. Coch. = cochlear ganglion.
Utric. = utricle.
Sacc. = saccule.
Sup., Post., Lat. = superior, posterior, and lateral semicircular canals.
Sac. end. = sacculus endolymphaticus.

from which they are separated by the perilymph; each canal commences as a dilation or *ampulla*, containing a projecting ridge, the *crista*. The canals open into the utricle by means of five apertures, one being common to the superior and posterior canals.

(ii) The *utricle* communicates with the *saccule* by means of the ductus endolymphaticus. Both the saccule and utricle contain a projecting ridge, the *macula*. The *canalis reuniens* unites the saccule and the duct of the cochlea.

Structure of Crista and Macula.—The crista and macula are the specific receptors of the vestibular apparatus and have a similar structure. Covering the ridge is a tall columnar epithelium (hair cells) giving attachment to long stiff *hairs* which project into a firm gelatinous material, the *cupula*

terminalis (Fig. 375). Between the hair cells lie the fibres of origin of the vestibular division of the eighth nerve.

(i) In the canals the cupula rises to the roof of the ampulla, acting as a movable partition which divides the ampulla into two compartments (Fig. 377).

(ii) In the saccule and utricle the cupula contains many chalky particles, the *otoliths*, hence the name the otolith organ. When the head in man is in the normal erect position, the macula of each utricle is approximately in the horizontal plane, with the cupula, hairs, and otoliths *rising vertically* from the macular epithelium; the macula in each saccule then lies in the vertical plane, with the hairs and otoliths projecting *horizontally* outwards into the cupula.

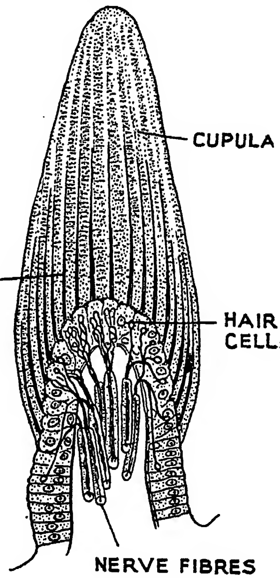


FIG. 375.—Structure of Crista of Ampulla of Semicircular Canal. (After Camis.)

Central Vestibular Connections.—The nerve endings in the maculæ and cristæ continue as nerve fibres which have their cell bodies in the bipolar cells of the *vestibular ganglion*; the central axons of the vestibular nerve enter the medulla ventral to the restiform body (inferior peduncle) and dorsal to the descending root of the Vth nerve. The axons divide into ascending and descending branches which end in four nuclear masses: (i) The large *medial* (principal) nucleus in the pons and medulla; (ii) the *descending* nucleus associated with the descending vestibular fibres; (iii) the *superior* nucleus (of Bechterew) at the level of the VIth nucleus; (iv) the *lateral* nucleus (of *Deiters*) in the lower pons. For all practical purposes these four nuclei can be treated as a single functioning entity. Some vestibular fibres pass *directly* to the *flocculonodular lobe* of the cerebellum (Fig. 382).

Fibres from the vestibular nuclei pass

(i) to the palæocerebellum of both sides via the restiform body (Fig. 383A).

(ii) directly and via the cerebellum to the red nucleus and the nuclei of the reticular formation in the brain stem (Fig. 373).

(iii) in the median longitudinal bundle to the oculomotor nuclei of both sides;

(iv) via the medial lemniscus to the opposite thalamus and thence to the opposite temporal lobe;

(v) down in the vestibulospinal tracts to the ventral columns of the spinal cord to end directly (and also via short interneurones) round ventral horn cells (Figs. 367, 368, 373).

Mode of Action of Otolith Organ (Saccule and Utricle).—The maculæ of the saccule and utricle are stretch receptors, the effective stimulus being the pull of gravity on the cupula and contained otoliths and hairlets;

the hair cells are thus deformed with resulting stimulation of the nerve filaments which lie between them. As might be expected the saccules are affected by a lateral tilt of the head: thus if the head is tilted laterally to the right (to rest on the shoulder) the cupula of the right saccule hangs downwards and pulls on its macula which is maximally stimulated; the cupula of the left saccule points upwards and " rests " on the macula, this being the position of minimal stimulation of the nerve endings. Ventral or dorsal flexion of the head (fore and aft tilt) affects the utricular maculæ; with the head erect the cupulæ in the utricles point upwards providing a minimal stimulus; when the head is bent well forward, or back, the cupulæ are pendent, pulling on the maculæ and so stimulating them maximally.

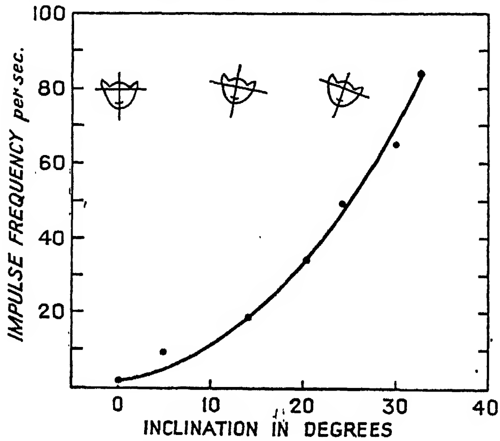


FIG. 376.—Response of Gravity Receptors (Otolith Organ). Relation between tilt of head and frequency of impulses in vestibular nerve. (After Adrian, *J. Physiol*, 1943, *101*, 393.)

Decerebrate cat. The degree of lateral tilt of the head is shown in the upper diagrams. The impulses were recorded from the *right* nerve while the head was being tilted to the *right* (*i.e.* right cheek down). As the tilt was increased the impulse frequency rose correspondingly. When the head was tilted to the *left*, the discharge in the *right* nerve ceased.

Action potentials can be recorded in the appropriate branches of the vestibular nerve or in its nucleus in the medulla. A tilt of as little as 2·5° stimulates the appropriate maculæ; as the tilt increases the frequency of the discharge progressively rises (Fig. 376). The general law of the receptors thus applies here, that the frequency of the impulse stream is directly related to the strength of stimulation. If the head is kept in any particular position, the impulse discharge pattern persists for as long as the position is maintained, except for some slight reduction in the discharge rate; the receptors thus show little adaptation (like the stretch receptors in muscle and tendon, p. 551).

Mode of Action of the Semicircular Canals.—An old experiment of Ewald illustrates how the semicircular canals are stimulated. Bore two holes in one bony canal (in a pigeon). Stop the one farthest away from the ampulla

with amalgam so as to block the membranous canal completely. Through the second hole introduce a rubber tube into the perilymph; when air is gently blown down this tube the membranous canal is compressed and the endolymph moves in the only direction it can do freely, namely, towards the ampulla. As a result nerve impulses are set up as judged by a resulting reflex movement of the head and eyes in the plane of the canal and in the direction of the endolymph current.

Direct observations have been made on the exposed semicircular canals in fish; a drop of oil is introduced into the canal, the fish is rotated, and the behaviour of the cupula is watched. As the rotation begins the endolymph in the canal moves, as shown by the shift in the position of the drop of oil; the cupula, which rises up as a septum completely dividing the ampulla in two, becomes bent over in the direction of the endolymph movement to an angle of up to 30° (Fig. 377).

The effective stimulus to each ampulla is rotation of the head in the plane of its canal. Consider the case of rotation of the head in the horizontal plane, in the direction of the arrows as shown in Fig. 378; the left ampulla is "leading" its canal while the right ampulla is "trailing" behind its canal. As the endolymph possesses *inertia* it does not move as fast, initially, as the canal in which it is contained; thus in a short-lived rotational movement of the head (*e.g.* one or two turns) the endolymph lags behind; this is equivalent to a flow in the *reverse* direction from that of the head movement. In the experiment illustrated by Fig. 378 the initial endolymph movement is thus towards the right ampulla and away from the left ampulla; both cupulæ presumably swing to the left. Action potentials led off from the appropriate nerves show that the frequency of the impulses from the right ampulla is increased while that from the left ampulla is decreased; *i.e.* in the case of the lateral (horizontal) canals, the "trailing" ampulla is stimulated while the "leading" ampulla is depressed.

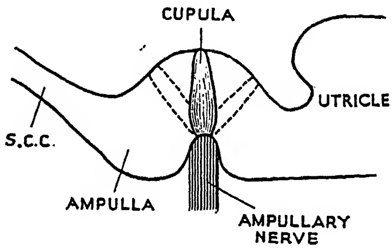


FIG. 377.—Mode of stimulation of Semicircular Canals.

S.C.C., semicircular canal.

The cupula, situated on the top of the crista, completely blocks the ampulla of the membranous canal. The cupula is caused to swing by movements of the endolymph.

The stimulus to the cristæ is obviously due to the swing of the cupula set up by the endolymph; it seems that a swing in a certain direction in any canal increases the stimulus to the nerve endings, while a swing in the opposite direction in that canal decreases the stimulus to the nerve endings. The combination of increased impulse discharge from one ampulla and decreased impulse discharge from the other, may form the basis of the interpretation of the direction of the movement. The degree of alteration of the frequency of the impulse discharge is directly related to the rate of acceleration of the rotational movement.

As the rotation is continued the endolymph takes up the same rate of movement as its canal; the cupula, by reason of its own elasticity, then returns (in about 30 seconds) to its original resting position, and the resting

nervous discharge is resumed (*i.e.* the discharge in the "active" ampulla decreases and that in the "depressed" ampulla increases). On cessation (or deceleration) of the movement, changes which are the reverse of the initial ones, occur. The endolymph, by reason of its momentum continues to move after the canal has come to rest; thus in Fig. 378 on cessation of head movement, the endolymph will continue to flow in the direction shown by the arrows, *i.e.* to the right; the cupulæ will swing to the right. As this is the opposite direction to that at the beginning of the movement the subjective impression will be of a movement in the reverse direction to the previous rotation. Finally the cupulæ by reason of their elasticity, regain their resting position, and the sense of movement ceases.

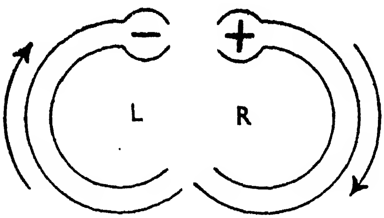


FIG. 378.—Diagram to illustrate Mode of Action of Lateral (Horizontal) Semicircular Canals.

The arrow shows the direction of movement of the head. The right ampulla is stimulated; the nervous discharge from the left ampulla is decreased.

The vertical posterior canal acts similarly except that for some unknown reason the ampulla which is *leading* is stimulated (Fig. 379, B); the details for the vertical superior canal are uncertain (Fig. 379, C). The left superior and the right posterior canals act as a functional pair, as do the right superior and the left inferior canals. With the head at rest there is a *steady "spontaneous" discharge* of impulses from all the six ampullæ. Characteristic modifications of this discharge pattern are set up by rotatory movements in any direction.

It must be repeated that the semicircular canals give information about *movements*, the otolith organ about the *position* of the head.

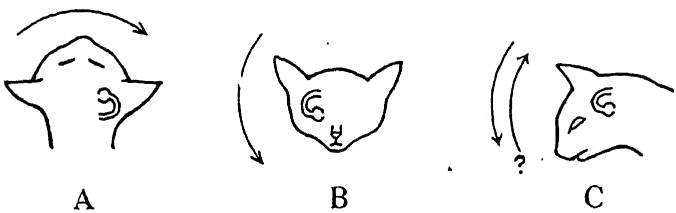


FIG. 379.—Direction of Rotation which Stimulates the Three Semicircular Canals (Cat). (Adrian, *J. Physiol.*, 1943, *101*, 397.)

A: horizontal canal; stimulating direction is with ampulla trailing.
B: posterior canal (rotation in transverse plane); stimulating direction with ampulla leading.
C: superior (anterior) canal (rotation in median plane); results uncertain.

EXTIRPATION EXPERIMENTS.—(1) *Bilateral* extirpation in birds produces grave disturbances; the animals subsequently cannot stand or fly or maintain any fixed position. In the cat or monkey the initial symptoms soon disappear and the animal behaves fairly normally if allowed the *use of its eyes*; it fails, however, to orientate itself under water and would soon drown; it cannot right itself when falling blindfold through the air (cf. p. 593). Muscle tone is not permanently decreased.

(2) *Unilateral* extirpation gives rise to complex derangements of postural activity.

(i) The *immediate* effects are skew deviation of the eyes (*i.e.* one eye rolls upwards and outwards and the other downwards and outwards), nystagmus, rotation and lateral flexion of the head (so that the occiput is turned to the side of the lesion); these changes are due to the unopposed action of the intact labyrinth. The altered position of the head sets up neck reflexes which, secondarily, modify the posture of the trunk and limbs. The limbs on the side of the lesion (the side to which the vertex is pointing, or skull (vertex) limbs) flex, and the limbs on the opposite side (jaw limbs) extend (p. 591). There is spiral rotation of the trunk.

(ii) The *permanent* effects are—

(*a*) Nystagmus: there is a slow swaying movement towards the side of the lesion and a quick return towards the midline.

(*b*) The reciprocal changes in tone and the head rotation persist.

(*c*) The rotation of the trunk diminishes.

(3) If one canal, *e.g.* the horizontal, is removed, spontaneous movements are set up in the plane of that canal, as no impulses are sent up to give information about the movement and thus check it.

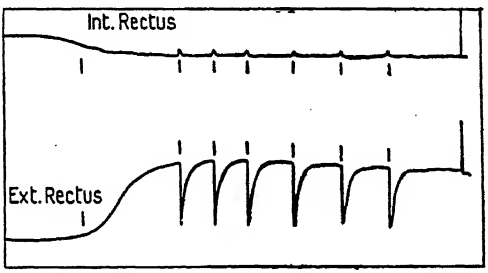


FIG. 380.—Reciprocal Innervation of Eye Muscles of Rabbit during Nystagmus set up by Irrigation of Ear.

Lower curve, external (lateral) rectus; upper curve, internal (medial) rectus. Decerebellate preparation. (After de Kleijn.)

RELATION OF VESTIBULE TO REGULATION OF POSTURE.—The experiment of Ewald, and the results of extirpation show that the vestibule normally plays an important part in the regulation of posture. Studies on the *tonic labyrinthine reflexes* (p. 591) and on the *righting reflexes* (p. 593), which have been fully considered, indicate how the otolith organ helps reflexly to maintain the upright position of the head and to adjust the position of the body to that of the head in space or relative to the trunk.

BARANY'S TEST.—The semicircular canals in man can be readily stimulated. By throwing the head backwards 60° and looking to the opposite side at an angle of 50° from the middle line and up to the ceiling, the *lateral* (horizontal) canal is placed *vertical*. The right ear, for example, is then syringed with cold air, which causes (by convection currents) a downward movement of the endolymph; this is equivalent to moving the head to the opposite side (to the left) (Fig. 380). The patient complains of giddiness, and if allowed to stand tends to fall to the right; nausea and vomiting may occur. The following results can be noted:

(*a*) *Nystagmus*.—The short jerk is to the opposite side (the left), and the slow movement is to the same side (the right).

The slow deviation to the same side is due to impulses which reach the eye nuclei from the vestibule via the vestibular nuclei. The short jerk is

probably a compensatory movement initiated by the cerebral cortex. The nystagmus illustrates well the law of reciprocal innervation (p. 545). De Kleijn removed the (right) eyeball in the rabbit and connected the internal and external rectus muscles to levers. Nystagmus was produced by syringing the (right) external meatus. The slow phase of the movement was to the right, and while the external rectus slowly contracted the internal rectus relaxed correspondingly slowly; in the quick phase the internal rectus contracted suddenly and the external rectus as rapidly relaxed (Fig. 380).[1] (Cf. p. 609.)

(β) *Past Pointing.*—On attempting to raise and lower the arms and touch a given point on a tape (when the eyes are shut), the limbs deviate out to the stimulated side. (Cf. p. 609.)

(γ) *Spontaneous deviation* of the limbs occurs towards the stimulated side.[2] (Cf. p. 609.)

Abnormal vestibular stimulation also produces complex *autonomic* disturbances such as alterations in blood pressure, heart rate, respiration and bowel tone and movements; these changes are well seen in the phenomena of *sea-sickness.*

THE CEREBELLUM[3]

Anatomy of the Cerebellum.—It is customary to divide the cerebellum into: (i) two large, laterally placed cerebellar hemispheres, and (ii) a small central portion, the vermis, so called " because it resembles a worm bent on

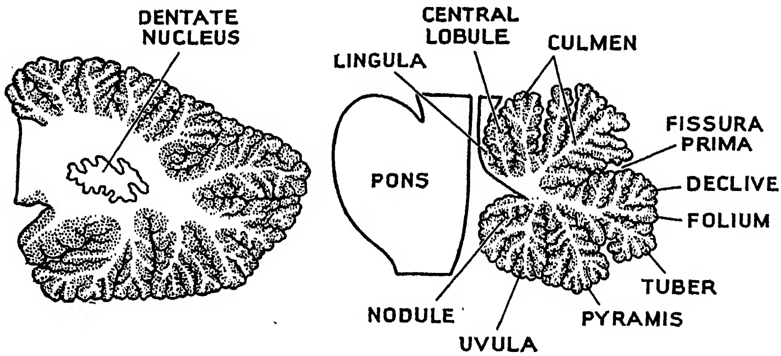


FIG. 381.—Anatomy of Cerebellum. (Modified from Ranson and Clark, *Anatomy of the Nervous System*, 1947, W. B. Saunders & Co.)

Sagittal sections through human cerebellum.
Left-hand figure: Section passes through lateral hemisphere and dentate nucleus.
Right-hand figure: Section passes through the vermis in the median plane.

---

[1] The clinical terminology of vestibular nystagmus is confusing, as attention is mainly paid to the " quick " or correcting component. Thus a so-called " right horizontal nystagmus " is one in which the quick movement is to the right, and the slow to the left. For completeness it may be mentioned that nystagmus may also be *vertical* or *rotary.*

[2] It will be noted that the tendency to fall, the spontaneous deviation, the past pointing, and the slow phase of the nystagmus are all towards the syringed side.

[3] Holmes, *Brain*, 1917, *40*, 461; *Lancet*, 1922, i, 1177 *et seq.* Symposium in *Brain*, 1927, *50*, 275. Assoc. Res. nerv. ment. Dis., *Cerebellum*, Baltimore, 1926. Dow, *Biol. Rev.*, 1942, *17*, 179. Adrian, *Brain*, 1943, *66*, 289. Fulton, *Functional Localization in the Frontal Lobes and Cerebellum*, Oxford, 1949.

itself to form almost a complete circle." The shape of the vermis is best seen in a sagittal section through its median plane as in Fig. 381. This conventional anatomical terminology does not, however, correspond to the functional differentiation or the phylogenic history of the cerebellum.

The older part (phylogenetically) of the cerebellum is called the *palæocerebellum* ; the newer part is called the *neocerebellum.*

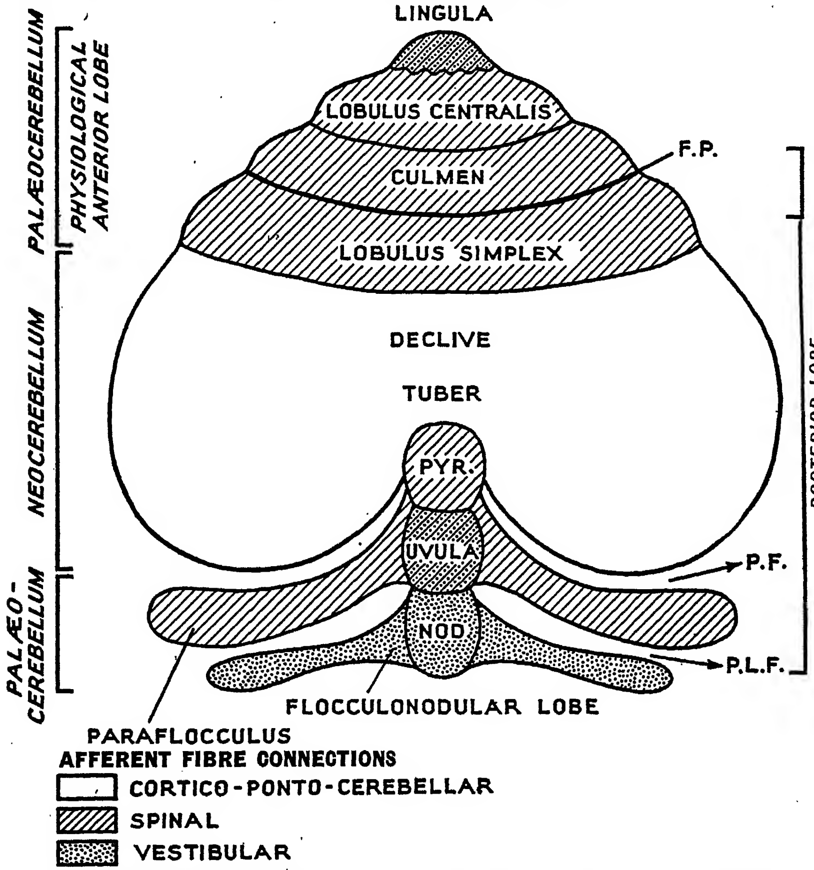


FIG. 382.—Diagram of Primate Cerebellar Cortex laid out flat and looked at from the Dorsal Surface to show Principal Subdivisions and Afferent Connections. (Modified from Dow and Fulton.)

F.P.=primary fissure, separating *anatomical* anterior obe from posterior lobe. (Physiological anterior lobe includes lobulus simplex.) P.F.=prepyramidal fissure. P.L.F.=postero-lateral fissure=posterior border of posterior lobe. *Anatomical* posterior lobe extends from P.F. to P.L.F.

(i) The palæocerebellum consists of the flocculonodular lobe (nodulus and the two lateral flocculi, the paraflocculi, most of the vermis (except the declive, folium and tuber) and the "physiological anterior lobe of the cerebellum" (lingula, lobulus centralis, culmen and lobulus simplex) (Fig. 382).

(ii) The neocerebellum consists of most of the *posterior lobe* ; it does not include the pyramis, uvula and paraflocculi or the lobulus simplex, but does include the folium and tuber (of the vermis) (Fig. 382). Roughly speaking the greater part of the lateral hemispheres of the cerebellum is neocerebellum.

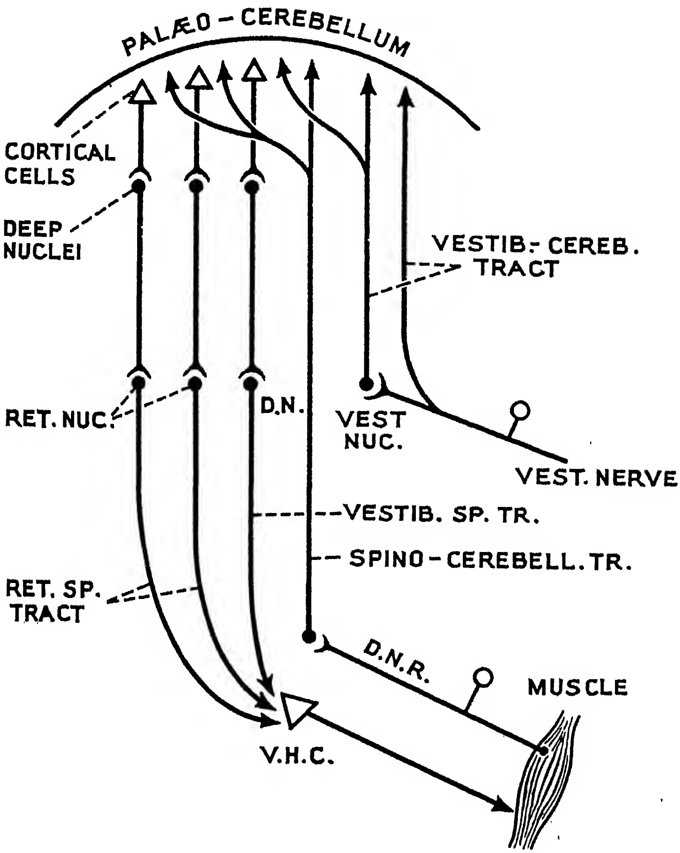


Fig. 383—Main Connections of Palæocerebellum.

D.N.R., dorsal nerve root.
V.H.C., Ventral horn cell.
Vest. Nerve, Vestibular nerve.
Vest. Nuc., Vestibular nucleus.
D.N., Deiters' nucleus.
Vestib-Cereb Tract, Vestibulo-cerebellar tracts.
Ret. Nuc., Reticular nuclei.
Ret. Sp. Tract, Reticulospinal tracts.
Spino-cerebell. Tr., Spinocerebellar tracts (dorsal and ventral).

The size of the neocerebellum increases phylogenetically with the development of the cerebral hemispheres.

As will be explained later, the palæocerebellum discharges to the brain stem and thus regulates posture. The neocerebellum discharges to the upper brain stem, thalamus and cerebral cortex and thus regulates voluntary movements as well as posture.

The surface of the cerebellum is covered by a grey cortex which is thrown into numerous fine convolutions (*folia*). In the white matter of the interior of the cerebellum are found masses of grey matter (*deep nuclei*) : the *dentate* nucleus in the lateral hemispheres, and a group of three more medially placed nuclei (*n. emboliformis*, *n. globosus* and *n. fastigii* (the last is part of the roof of the fourth ventricle and therefore referred to as a " roof nucleus ")).

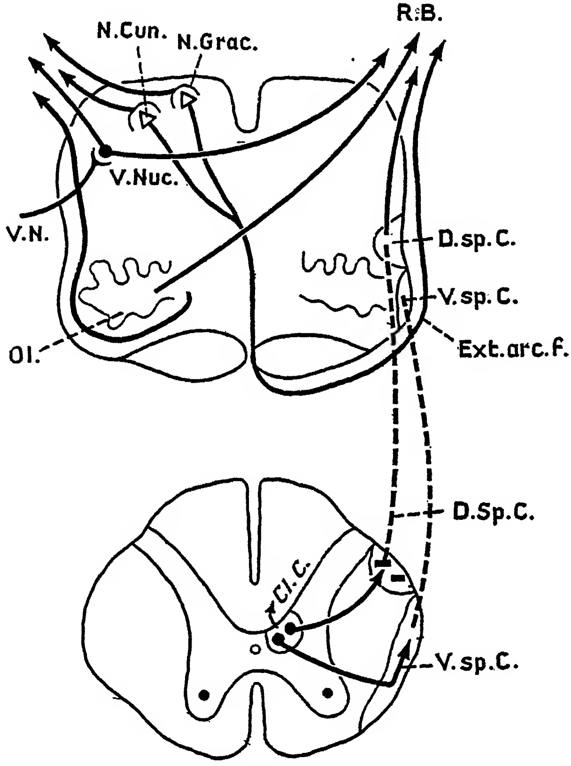


FIG. 383A.—Diagram of Spinal Cord and Medulla to show Ascending Cerebellar Tracts.

V.sp.C., ventral spinocerebellar tract. D.sp.C., dorsal spinocerebellar tract. Cl.C., Clarke's column of cells. Ol., Olivary nucleus. V.N., vestibular nerve. V.Nuc., vestibular nuclei. N.Grac., nucleus gracilis. N.Cun., nucleus cuneatus. R.B., restiform body. Ext.arc.f., external arcuate fibres.

Histology of the Cerebellar Cortex.—Three layers are recognized—an outer *molecular*, an inner *nuclear*, and a middle layer of *Purkinje* cells.

(1) The *Purkinje* cells are large flask-shaped cells ; their axons pass into the white matter to end mainly in the *deep cerebellar nuclei* (*supra*). Their dendrites form a complex tree-like arborization in the molecular layer.

(2) The molecular layer contains the *basket cells*, whose axons link up a number of Purkinje cells and end in basket-like interlacements round their bodies.

(3) The nuclear layer consists of round cells (*granules*) with large nuclei which practically fill the cell substance. Their axons pass up into the molecular layer and bifurcate, but their mode of termination is not known. Their dendrites end locally in their own layer in tuft-like endings.

(4) Afferent fibres enter from the white matter: the *moss* fibres end in tufts round the dendrites of the granule cells; the *tendril* or climbing fibres pass into the molecular layer and twine round the dendrites of the Purkinje cells.

The cerebellar cortex is thus an elaborate internuncial mass connecting the incoming afferents with the *deep nuclei*; it is the latter (in the main) which give rise to the efferent cerebellar fibres which pass to the brain stem, thalamus, and cerebral cortex.

Connections and Functions of the Cerebellum. 1. FLOCCULONODULAR LOBE.—(i) Its main *connections*, afferent and efferent, are with the *vestibular* nuclei (Fig. 382).

Vestibular afferents (from the otolith organ and semicircular canals) pass both directly or after relaying in the vestibular nuclei via the restiform body (inferior peduncle) to the flocculonodular lobe; efferents from this lobe return in the restiform body to the vestibular nuclei. From these nuclei the descending vestibulospinal tract connects with spinal motor neurones. The flocculonodular lobe is thus a long relay, superimposed *on the vestibular nuclei, for controlling bodily posture.*

(ii) *Effects of Extirpation.*—Isolated extirpation of this lobe has no effect on voluntary movement or on limb posture. The main disturbance is of *equilibrium,* shown by an inability to maintain the *erect position.* A monkey so afflicted is unable to stand up without swaying and generally sits in a corner propped up by the two sides of its cage. There is a common cerebellar tumour (*medulloblastoma*) occurring in childhood which, because it affects the flocculonodular lobe, produces similar unsteadiness in standing and walking, the so-called *trunk ataxy.* Similar symptoms (as might be expected) follow section of the restiform body.

After removal of the *nodulus* in the dog, the animal no longer develops motion sickness.

2. **Main Palæocerebellum.**—(1) CONNECTIONS (Figs. 382, 383).

(i) *Afferent Fibres.*—The main part of the palæocerebellum, *i.e.* the "physiological anterior lobe" and the posterior part of the posterior lobe, receives afferents chiefly from the *muscles and related deep structures,* and to a minor extent from the vestibular apparatus.

The details are as follows (Fig. 383A):

(*a*) Muscle afferents from the trunk and limbs enter the dorsal nerve roots and end in Clarke's column at the base of the dorsal horn of grey matter. The axons arising from these cells form:

(α) The dorsal spinocerebellar tract which passes up the lateral column of its own side (in its dorsal part near the surface) to enter the restiform body (inferior peduncle of the cerebellum).

(β) The ventral spinocerebellar tract which passes up the ventral column of its own and also the opposite side, to enter the brachium conjunctivum (superior peduncle of the cerebellum).

(*b*) Muscle afferents from the trunk and limbs also pass up in the dorsal columns, relay in the nuclei gracilis and cuneatus in the medulla and then pass via the restiform body to the cerebellum.

(*c*) Muscle afferents from the head, end in various sensory nuclei (especially V) and pass thence to the cerebellum.

(ii) *Efferent Fibres.*—Efferents from this part of the palæocerebellum pass via the deep nuclei to the brain stem (mainly in the restiform body and the brachium pontis (middle peduncle of cerebellum)) to end in the vestibular nuclei and the reticular nuclei of the pons and medulla. The descending vestibulospinal and reticulospinal fibres connect with the lower motor neurones; ascending fibres from the vestibular nuclei (in the median longitudinal bundle) end round cranial nuclei especially those supplying the eye muscles. The main palæocerebellum is *thus a long relay superimposed on the brain stem and* (*as shown below*) *regulating the posture of the eyeballs, head, trunk and limbs* (Fig. 383).

(2) Functions.—Most of the experimental work has been carried out on the anterior lobe. As explained on p. 593 some of the descending tracts arising in the brain stem and ending at the lower motor neurones are facilitatory (increasing muscle tone); others are inhibitory (decreasing tone). It has been shown that the anterior lobe, by controlling the cells of origin of these tracts can modify muscle tone in complex ways.

(i) Stimulation of the anterior lobe may *inhibit* tone in the extensor muscles both in the intact and in the decerebrate animal; the effect is most marked on the same side of the body. The inhibitory impulses pass via the *dentate nucleus* to the brain stem. Conversely, ablation of the anterior lobe enhances extensor tone in decerebrate rigidity. Extirpation of the lingula (which receives vestibular afferents (Fig. 382)) leads to enhancement of vestibular reflexes.

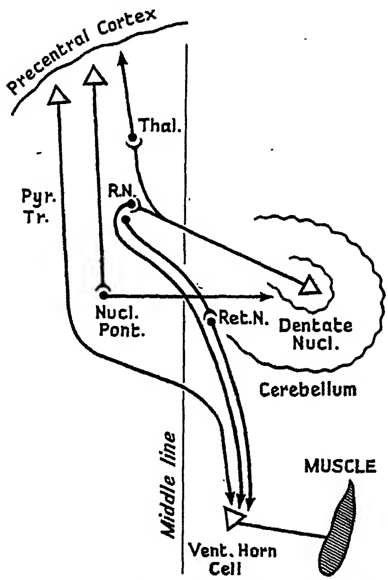


Fig. 384.—Main Connections of Neocerebellum.

Thal., Thalamus. Pyr. Tr., Pyramidal tract. R.N., Red nucleus. Nucl. Pont., Nuclei pontis. Ret.N., Reticular nuclei.

(ii) On the other hand stimulation of the anterior lobe may sometimes *increase* muscle tone; the facilitatory impulses pass via the *n. fastigii* to the brain stem.

(iii) The effects of the anterior lobe on the lower motor neurones are to some extent topographically localized. Specific areas in the anterior lobe facilitate or inhibit the discharge of specific groups of motor neurones. Electrical studies have shown that impulses from specific parts of the limbs project on to specific points in the anterior lobe.

The rôle of the palæocerebellum as a reflex centre for the finer control of posture has been summarized by Sherrington: "the cerebellum may indeed be described as the head ganglion of the proprioceptive system and the head ganglion here, as in other systems, is the main ganglion." [1]

[1] By "head ganglion," Sherrington means the ganglion situated in the head.

3. **Neocerebellum** (Most of the Cerebellar Hemispheres).—(1) CONNECTIONS (Fig. 384).—The neocerebellum also receives some proprioceptor afferents from muscles via the spinocerebellar tracts. The main connections of each cerebellar hemisphere, however, are with the *opposite cerebral hemisphere* and *the upper brain stem.*

(i) *Cortico-pontine fibres* arise in the frontal lobes mainly in the excitomotor cortex (areas 4 and 6) and in the temporal lobes, and end in the nuclei pontis, which also receives fibres from the pyramidal tracts. From the nuclei pontis the *pontocerebellar fibres* arise, which cross to the opposite cerebellar hemisphere in the brachium pontis.

(ii) (*a*) From the *dentate nucleus* of the lateral lobe of the cerebellum, the *brachium conjunctivum* [*superior cerebellar peduncle*] arises. It crosses the middle line in the midbrain ; most of the fibres end in the (opposite) *thalamus* [latero-ventral nucleus]. A new relay arises there to end in the (opposite) cerebral hemisphere, mainly in the excitomotor cortex (areas 4 and 6). The pathway, cerebellum–thalamus–cerebral cortex–pons–cerebellum, can be regarded as a closed circuit by means of which the cerebrum and cerebellum mutually influence one another's activities.

Each cerebellar hemisphere controls the opposite cerebral cortex ; in its turn the excitomotor cortex via the pyramidal tract controls the movements of the opposite side of the body. Because of the double decussation (*i.e.* of the brachia conjunctiva [superior cerebellar peduncles] and of the pyramidal tracts), *each cerebellar hemisphere controls voluntary movements on its own side of the body.*

(*b*) Some of the efferent cerebellar fibres end in the opposite *red nucleus* and in the *nuclei of the reticular formation* of the brain stem. Descending rubrospinal, rubroreticular and reticulospinal fibres end at the lower motor neurones and influence tone.

The connections of the neocerebellum shown in Fig. 384 make it clear how it can regulate voluntary movement and influence posture.

(iii) Electrical stimulation of the motor area of the cerebral cortex produces localized responses in the cortex of the opposite cerebellar hemisphere ; face, arm, and leg areas in the motor cortex link up with specific areas in the cerebellar cortex. Conversely, cerebellar stimulation modifies the " resting " action potentials in the cerebral cortex (area 4) and enhances the excitability of the motor cortex to artificial stimulation.

(iv) The anatomical and functional relationships between the palæocerebellum and neocerebellum are obscure. There is no evidence that the main palæocerebellum and the neocerebellum are discrete physiological entities. It is difficult to understand how the neocerebellum can function effectively with so sparse a supply of afferents from the muscles and the vestibule. There may possibly be important connections between the different areas of the *grey surface* of the cerebellum which would convert the organ into an integrated unit.

(2) **Results of Lesions of Neocerebellum.**—Injuries to the cerebellar hemisphere in man (due to gunshot wounds or tumour) produce characteristic disturbances of posture and of voluntary movement. In unilateral lesions, the changes are mainly found on the same side of the body.

The detailed findings are set out fully below.

Disturbances of Posture.—(1) ATONIA.—There is diminution or loss

of tone in the muscles of the same side of the body, particularly in the limbs The muscles feel flabby and are readily compressed and displaced; they can be stretched without producing discomfort. If the forearm is held and the hand shaken to and fro, the oscillations of the hand are normally limited owing to the tension of the muscles of the forearm. In the patient the hand swings about like a flail, inertly, till the joints lock and the bony structures prevent further movement. The limbs tend to assume or to be placed in abnormal positions that would normally be avoided. If the patient sits up in bed, the weight is thrown on the *dorsum* of the affected hand, so that the fingers and wrist are overflexed to a degree that would ordinarily be painful. Normally, if the forearm be held up vertically, the hand is only slightly flexed. In a case with cerebellar injury, extensor tone is lost, and the hand falls passively into a position of extreme flexion. These symptoms illustrate the varied disturbances which may follow loss of postural activity, and indicate how the case should be examined clinically.

(2) Attitude.—The face is usually rotated towards the opposite side, and the occiput is approximated to the affected shoulder. The homolateral shoulder is slightly raised and is in front of its fellow of the other side. The leg is abducted and rotated outwards, thus giving the body a wider basis of support. The weight is thrown on the sound leg, and so the trunk is bent with the concavity towards the affected side. Giddiness (vertigo) is uncommon, except immediately after the injury, and when present has no constant relation to the side of the injury. In cases of tumour, external objects appear to move away from the side of the lesion, and the body tends to follow suit, the head and shoulder, as mentioned above, turning away from the affected side.

(3) Spontaneous Deviation.—If the eyes are closed and the arms are held straight out in front of the body, they normally remain quite steadily in this position. In cerebellar disease, the homolateral arm sways slowly or quickly outwards, away from the symmetrical position, and then comes to rest.

(4) Barany's Pointing Test.—The patient lies in bed, and a tape is held above the bed; he is asked to touch a certain spot on the tape, then close the eyes, bring the finger down to the bed and then raise it again to the original spot on the tape. As the movement is repeated, the finger gradually deviates outwards. Both this sign and spontaneous deviation are not due to loss of muscle sense, because the patient is aware of the position of the deviating finger and can touch it with his normal finger. They result from some disturbance of postural activity.

(5) Static Tremor.—There may be slight oscillations of the head and trunk in any direction, owing to irregular contraction of the muscles which maintain attitude. If the arm is held out, a tremor develops after some time which consists of a slow downward and a quick upward movement. The slow displacement is possibly due to the action of gravity on the tired limb, and the quick jerk may be the voluntary effort to get the arm back to its original position.

(6) Nystagmus.—Normally the postural mechanism keeps the eyes when at rest in the central position; when an object is being looked at the gaze is a steady one. In cerebellar lesions the eyes tend to deviate from the central position, some 10° to 30° to the opposite side. Why the deviation is in this direction is not clear, but having gained this position the eyes come to rest

and the position is called the *rest point.* On looking fixedly at an object elsewhere in the visual fields, especially towards the affected side, the eyes display jerking movements termed *nystagmus.* In character they resemble the static tremor described. There is a slow swaying movement towards the rest point, and a quick abrupt recovery, which is of cortical origin. In *tumour* of the lateral lobe the details of the nystagmus are a little different. The slow swaying movement occurs on looking to the *side* of the lesion, and there is a rapid to-and-fro jerking on looking to the opposite side.

(7) Alteration of the Deep Reflexes.—In cerebellar lesions, the jerks may be slower than normal and perhaps less vigorous as well, because the extent of the contraction produced by stretch depends on the initial degree of tone in the muscle, and in cerebellar disease tone is diminished; nor is there any tendency to maintain the quadriceps in the shortened position, so that the limb falls quite passively when the twitch passes off. If the patient is seated on a high stool, it is found that the leg on falling goes on swinging freely to and fro. This type of response is therefore called the *pendulum knee-jerk* (Fig. 385).

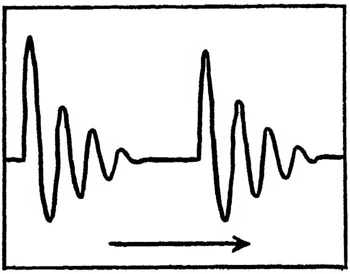


Fig. 385.—Cerebellar Knee-jerk.

Two responses are shown. Note in each case the rapid relaxation and the after-swinging (pendulum knee-jerk). (Gordon Holmes, *Lancet,* 1922.)

Disturbances of Voluntary Movement.—(1) There is feebleness (*asthenia*) of moderate degree, especially in the arm. This becomes more obvious when the movement is prolonged, or requires exertion of much power. The muscles tire very readily; the arm tends to droop if held outstretched for any time. There is *slowness* in carrying out every phase of a movement—delay in initiating, in attaining full power, in commencing relaxation, and in reaching full relaxation. There is reluctance to move the affected limb, which lies for long periods in one position. Objects are always picked up with the sound hand; the affected arm *does not swing* to and fro in walking.

(2) *Ataxia* is present—*i.e.* clumsiness and disorderliness of movement. Patients say: "My hand will not go straight, it is drunk; I do not seem to be able to do what I want with my hand, though if I get hold of anything I can grip it all right." The ataxy is worse if a movement is hurried or if a number of joints are involved. It is not affected by closing the eyes, because muscle sense is normal and knowledge of position is unimpaired. The ataxy has been analysed thus:

(i) *Decomposition of the Movement.*—There is inability to perform simultaneously the various movements which constitute an act. The separate movements are carried out "by numbers" as in a gymnasium. Thus, on trying to bring the heel on the opposite knee while lying in bed, the heel is dragged along the bed till it reaches the knee, and then raised to the proper height, then adducted and placed on the knee.

(ii) *Asynergia.*—Proper co-ordination is lacking between the protagonists, the antagonists, and the synergic muscles which fix joints. Thus the extensors of the wrist are normally synergic muscles to the flexors of the fingers, the former placing the wrist-joint in the most favourable position for the

action of the latter (cf. p. 651). This can be readily seen on attempting to flex the fingers with the wrist horizontal or flexed. In cerebellar trouble, on flexing the fingers the wrist is extended too late or too early, or not at all.

(iii) *Dysmetria.*—The force employed is ill-adapted to the aim of the movement. The hand shoots past an object or is stopped prematurely.

The ataxia is attributed to lack of cerebellar "guidance" of the centres in the cerebral cortex concerned with voluntary movement. The synergist component of a voluntary movement is imperfectly employed (asynergia); there is disturbance of the duration of the contraction, of the proper timing of one concentration relative to the next (decomposition), and of the proper force of contraction (dysmetria).

These disturbances can be demonstrated by simple clinical tests:

(i) *Finger-Nose Test.*—The patient is asked to bring the tip of the finger of the outstretched hand on to the tip of the nose with the eyes shut. The movement is not begun promptly, the different groups of muscles do not co-ordinate efficiently, the shoulder-joint is not kept still, and the force is not nicely adjusted. The finger thus advances by a series of irregular jerks, the movement being broken up and each phase overdone; the finger shoots past the nose, or strikes the nose with excessive force and "keeps on rubbing itself in."

(ii) *Adiadochokinesis.*—The patient is ordered to carry out rapid pronation and supination movements simultaneously in both forearms. The movements on the affected side are slower, less regular, and the range is less uniform; adventitious movements occur at other joints owing to contraction of unnecessary muscles. Because of fatigue and disinclination the movements are soon abandoned (Fig. 386).

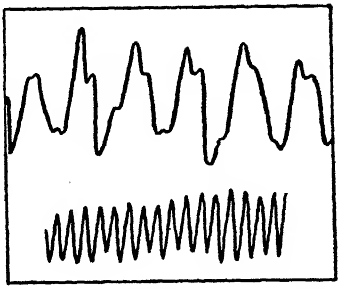


Fig. 386.—Cerebellar Ataxy. Adiadochokinesis.

Records of Rapid Pronation and Supination of Arm.

Upper curve—on side of cerebellar lesion (showing adiadochokinesis); lower curve —normal side. (Gordon Holmes, *Lancet*, 1922.)

(3) Gait.—On first getting out of bed, the whole body sways irregularly, and there is danger of falling to the affected side. The patient feels as if an invisible hand were pulling him in that direction. After a time, he can maintain his equilibrium fairly well and makes appropriate though clumsy movements to regain his balance if it be threatened. The gait is frequently described as resembling that of a drunken man. That is inaccurate; the patient walks carefully; he does not trust the affected side, taking short steps with that leg and hurrying off it. He *deviates spontaneously towards the affected side,* and then tries to bring himself back to the original line, thus taking a zigzag path.

(4) Speech.—This is affected in the same way as any other complex movement. It may be slow and monotonous, staccato or scanning in character; the consonants are frequently blurred, and there is a tendency to explosive utterance. Great efforts are made to utter any sentence, and these are associated with excessive facial grimacing.

After a time, *compensation* occurs, and the movements improve in efficiency.

There is less deviation, less decomposition, and better synergia. Excessive force and range persist, however, so that the movements remain jerky in character.

Clinical Conditions associated with Cerebellar Dysfunction.—

(1) In acute irritative lesions (*e.g.* vascular lesions) of the cerebellum, giddiness is severe and forced movements occur, which turn the patient so that the face on the side of the lesion is in contact with the pillow.

(2) *Tumour* of the neocerebellum produces signs closely resembling those already detailed. The nystagmus, as stated, consists of a slow to-and-fro movement on looking to the affected side, and a rapid to-and-fro movement on looking to the opposite side. If the flocculonodular lobe is involved, the signs are bilateral and mainly involve the *trunk* (p. 606).

(3) Tumours growing from the sheath of the *eighth* nerve usually involve the cerebellum later in their course.

(4) In *disseminated sclerosis*, plaques of the disease frequently occur in the cerebello-pontine region, and produce signs of cerebellar dysfunction such as nystagmus, intention tremor, and ataxy of speech and of other movements.

(5) In a group of diseases called *hereditary ataxy*, of which Friedreich's disease is the best known, the spinocerebellar tracts or other cerebellar connections tend to degenerate early, producing characteristic signs.

(6) Lesions of the cord, *e.g.* syringomyelia, or tumour, may press on the spinocerebellar tracts and cause nystagmus, etc.

STRUCTURE OF CEREBRAL CORTEX.[1] METHODS OF DETERMINING CORTICAL CONNECTIONS. THE ELECTROENCEPHALOGRAM

Structure of Cerebral Cortex.—Histological studies carried out over many years have led to the cerebral cortex being subdivided into many areas each with its distinctive cellular arrangement. The detailed cell structure of any cortical area is called its *cytoarchitectonics* or its *cytoarchitecture*. It is hoped that the results obtained, apart from their intrinsic interest, may help to solve problems connected with the localization of function in the cortex, as it is believed by some that each cortical area with a distinctive structure is a functionally specialized zone. This view may be correct about some cortical regions, *e.g.* the visual area, but it does not hold for all cortical areas. For example the two areas in the precentral motor cortex labelled areas 4 and 6 probably subserve some functions in common. Again the postcentral gyrus is subdivided into areas called 3, 2 and 1 which may subserve a single common function. It must always be remembered that the different cortical areas are *closely knit together* as will be described later (p. 616).

According to Economo, in typical regions of the cortex, *six* cell layers can be recognized which are numbered I to VI from without inwards (Fig. 387).

LAYER I: *Molecular layer*, which contains numerous dendrites, axons, and glia cells.

[1] Economo, *Cytoarchitectonics of the Human Cerebral Cortex*, London, 1929. McCulloch, *Physiol. Rev.*, 1944, *24*, 390; Res. Publ. Assoc. nerv. ment. Dis., *Frontal Lobes*, 1948.

Layer II : *External granule layer*, consisting of small densely packed cells (*granules*) which are round, polygonal or triangular in shape.

Layer III : *External pyramidal layer* which contains large pyramidal cells, usually increasing in size from without inwards.

Layer IV : *Internal granule layer*, resembling layer II.

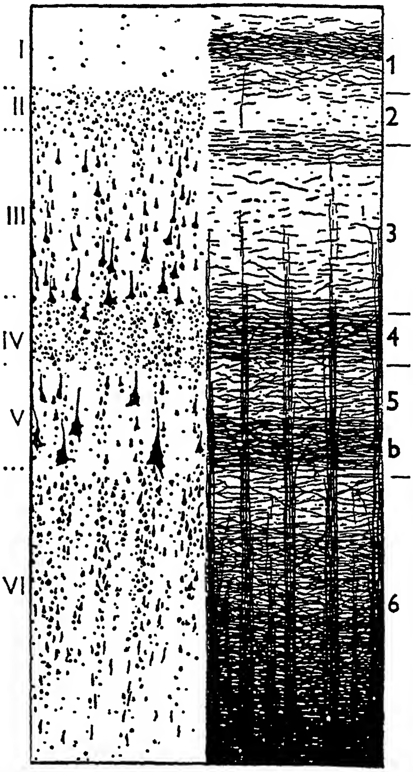


Fig. 387.—Diagram of the Layers of the Human Cerebral Cortex. On the left, I to VI, are the cell layers; on the right, 1 to 6, are the fibre layers as they appear in specimens stained by the Weigert-Pal method. Layers 4 and 5b are respectively the external and internal lines of Baillarger. (Economo, *Cytoarchitectonics of the Human Cerebral Cortex*, Oxford University Press, 1929.)

Layer V: *Internal pyramidal layer*, consisting typically of cells which resemble those found in layer III.

Layer VI : *Fusiform cell layer*, consisting of long spindle-shaped cells arranged perpendicularly to the surface.

The arrangement of the nerve fibres in the cortex is shown diagrammatically in Fig. 387. There are three principal bands of transversely running fibres in layers 1, 4 (outer line of Baillarger), and 5b (inner line of Baillarger), respectively. The longitudinally running fibres penetrate outwards as far as layer 2.

In addition to the three types of cells already referred to, *i.e.* the granule, pyramidal, and fusiform cell, there are special cells in certain regions, *e.g.* giant cells of Betz (60–120 μ by 30–80 μ) in area 4γ in the precentral gyrus and the giant stellate cells in the visual cortex.

Apart from the typical cortex just described (or *isocortex*) there is the *allocortex* which is constructed on an entirely different plan. The allocortex includes the uncus, hippocampus, and the gyrus dentatus; in man it constitutes about one-twelfth of the cerebral surface. In some species, however, for example the hedgehog, it may form about three-fourths of the surface.

The cortical grey matter differs in thickness in different regions, *e.g.* on an average on the convexity it is 3·5 mm., at the base 3 mm., and on the medial surface 2·7 mm. The maximum thickness (4·5 mm.) is found in the precentral gyrus and the anterior part of the temporal lobe; the minimum

(1·3 mm.) is at the frontal and occipital poles. The grey matter is twice as thick on the crown of a gyrus as in the furrow.

According to Economo, five fundamental types of isocortex may be differentiated; types 2, 3, and 4 are essentially alike and differ one from the other in details; types 1 and 5, on the other hand, contain very obvious distinctive features (Fig. 388).

Types 2, 3, *and* 4.—These have the six typical laminæ previously described.

(i) *Type* 2.—*Frontal* type (anterior two-thirds of the frontal lobe, superior

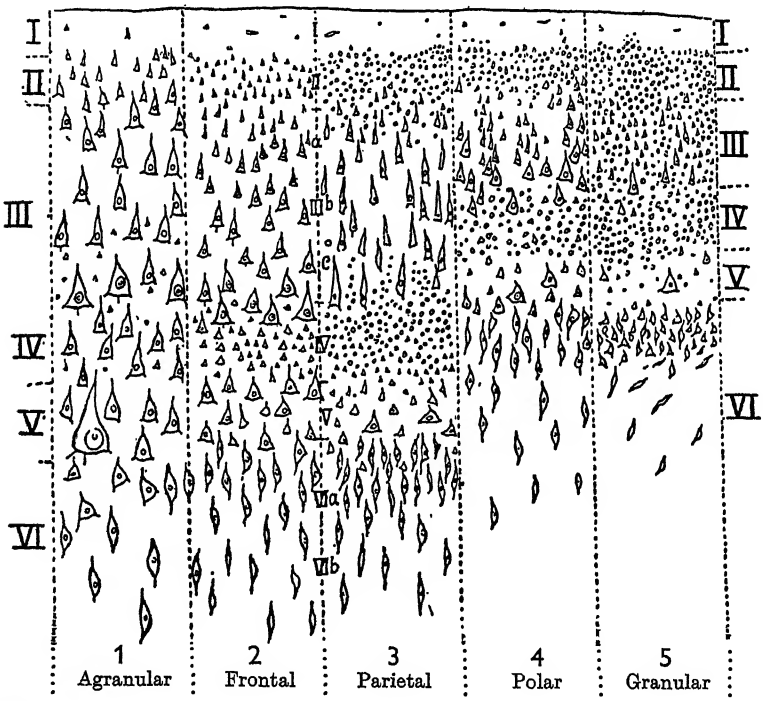


Fig. 388.—Diagram of the Five Fundamental Types of Structure found in the Cerebral Cortex. (Economo, *Cytoarchitectonics of the Human Cerebral Cortex*, 1929.)

parietal lobule, and part of the temporal lobe). The granule cells are triangular.

(ii) *Type* 3.—*Parietal* type (parietal lobe and junctional region of parietal, occipital, and temporal lobes). There is an increase in the depth and density of the two granule layers II and IV, and these cells are round in shape; the pyramidal cells are smaller, slender, and more numerous.

(iii) *Type* 4.—*Polar* type (only at the frontal and occipital poles). The cortex is narrow, and all the layers are reduced in depth though the cells are more densely packed.

Type 1.—*Agranular* cortex.—This is characteristic of the *excitomotor* regions of the cortex and is thus found in the posterior third of the frontal

lobe anterior to the fissure of Rolando. It is however also found on the convexity and medial surface of this region, in Broca's area, and the anterior part of the island of Reil. As the name implies, granule cells are completely *absent*; the cells in layer II and IV have become *pyramidalized, i.e.* replaced by pyramidal cells.

Type 5.—*Granular* cortex.—This is characteristic of the *sensory* zones, *e.g.* the postcentral gyrus (general body sensation), calcarine region (vision), Heschl's gyrus (hearing), The granules have largely replaced the pyramidal cells in layers III and V, *i.e.* the cells have become *granulized.* In the visual cortex the internal granule layer IV is divided into two parts by transversely running fibres (the line of Gennari).

Maps illustrating some of the cytoarchitectonically discrete areas are shown in Figs. 407, 430 *et seq.*; such maps are useful when stimulation or extirpation experiments are planned to determine more precisely the functional attributes of different cortical regions.[1]

Methods of Determining Cortical Connections.—(1) USE OF WALLERIAN DEGENERATION AND CHROMATOLYSIS.—(i) If the cell of origin of an axon is destroyed, the axon undergoes degeneration throughout its course; the path taken by the degenerating axon from its origin to its termination can be followed using the Marchi staining technique (p. 499). Thus if the "arm" area of the motor cortex is excised, the degenerating arm fibres of the pyramidal tracts can be traced through the central nervous system to end in the lower cervical and upper thoracic cord (Fig. 412). Similarly, if a nucleus in the thalamus is destroyed, the axons from the nucleus can be traced to their terminations in the cerebral cortex.

(ii) If an axon is cut, its cell of origin undergoes chromatolysis (p. 494). Thus, if the corticospinal fibres are cut in the medullary pyramid, their cells of origin in the motor area show chromatolysis.

(2) ELECTRICAL METHODS.—(i) As the nerve impulse is accompanied and signalled by a spike potential, the route and site of termination of nerve fibres can be demonstrated by electrical means. Thus by exploring various points in the central nervous system with micro-electrodes, when a certain tract is transmitting impulses, the course followed by the tract can be determined. This method has been used to trace, *e.g.* the course of the central auditory pathway (in the lateral lemniscus), and that of the pyramidal tract.

(ii) The points of termination of impulses from various sense organs in the cerebral cortex can, similarly, be determined. A minute electrode, placed on the surface of the cerebral cortex, picks up the spike potentials developed in its neighbourhood. A sense organ (eye, ear, skin receptor) is stimulated; the impulses set up finally reach the related "receiving" area in the cerebral

[1] It must be remembered that, as the necessary histological work is extremely difficult and time-consuming, very few brains in any one species have so far been exhaustively studied. As many cytological features must be taken simultaneously into account in delimiting an area, the results reported by different investigators show significant differences. The margins of the described areas are not sharply defined, but a gradual transformation seems to occur. Most workers use the map prepared by Brodmann which must, however, be regarded as merely a first approximation which is undergoing progressive modification as experience grows.

For a critical appraisal of the value of the results of cytoarchitectonic studies see Walshe, *Critical Studies in Neurology*, 1948. Walshe goes so far as to suggest that some of the details appearing in cytoarchitectonic maps may constitute contributions to neuromythology rather than to neurology.

cortex where their arrival can be detected by the occurrence of a sudden burst of electrical activity. By such means we can determine in detail how any sensory field is projected on to the cortex (p. 568).

(iii) PHYSIOLOGICAL NEURONOGRAPHY.[1]—This method is a very valuable modification of the electrical technique just considered. A piece of blotting-paper a few square millimetres in area is soaked in a strychnine solution and applied to the surface of the cerebral cortex; the strychnine stimulates the nerve cells locally, causing them to discharge; the nerve impulses, so generated, travel along the axons into the white matter to reach another area of the cortex or some subcortical nucleus. Let us suppose that we are dealing with nerve fibres arising in cells in region A in Fig. 389 and passing into the white matter to re-enter another region of grey matter, B, thus establishing connection with it. An electrode placed over B detects the arrival of the nerve impulses by recording the associated electrical disturbances (the spike potentials). Fig. 390 illustrates an experiment in which application of strychnine to the area shown as a black square gave rise to spikes in the points labelled respectively *a* and *b*, proving that fibres from the stimulated zone made connections with the latter regions. This potent technique has revealed an extremely complex system of interconnections in the cerebral cortex: (i) within areas which are cytoarchitectonically discrete; (ii) between discrete areas on the same side; (iii) between such discrete areas and corresponding points in the opposite cerebral cortex via the corpus callosum; (iv) between specific regions of the cerebral cortex and various subcortical ganglia. Some of the connections demonstrated by this technique are illustrated in Fig. 390, which should be consulted. The fact that *corresponding* points in the two hemispheres of the cerebral cortex are linked together is of great importance because it enables functionally identical areas in the two hemispheres to act as a physiological unit.

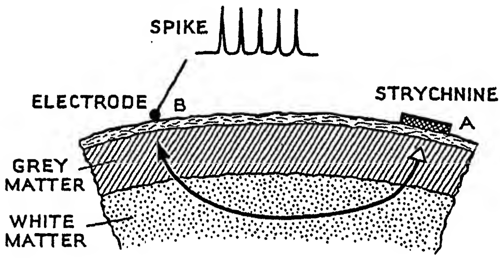


FIG. 389.—Diagram illustrating the Principle of the Method of Physiological Neuronography.

Cells give rise to axons, which run in the underlying white matter to re-enter the grey matter and there end. Application of strychnine at A to the cells stimulates them; impulses pass along the axons to the connected region of cortex B, where they can be recorded as "spikes" by a locally placed electrode.

The Electroencephalogram.—The electroencephalogram is the record obtained using indirect leads, of the changes in electric potential taking place n the grey matter of the cerebral cortex; the electrodes are placed on two points on the surface of the skull. A typical human record is shown in Fig. 391. Most of the waves normally recur at a rate of 9 or 10 per second (*alpha* waves); their amplitude is about 50 microvolts. This kind of record does not represent the activity of a single cell or of a few cells, but is the algebraic summation of the spike potentials generated by a *mass* of closely packed cells.

[1] For review see McCulloch, in Bucy, *Precentral Motor Cortex*, University of Illinois, 2nd edn., 1950.

The only circumstances under which the summed potentials, led off from a *mass* of cells, could appear as fairly regularly recurring potential waves, would be when many cells were discharging approximately *synchronously*, at the *same* rate; if the cells were discharging asynchronously and at *different* rates, the resulting record would consist of innumerable irregular excursions.

The large regular alpha waves appear only in the *absence* of mental and especially of visual activity; the eyes must be closed and the mind must be at rest; as soon as the eyes are opened, the alpha waves are replaced by very small irregular oscillations (Fig. 391). After the subject has been in the dark for some minutes, the alpha waves may persist even when the eyes are open, so long as *nothing is being looked at*; but if some invisible object is being fixated in the dark, the alpha waves disappear. It is clear from these results that the cortical cells, when not engaged in any specific task (*i.e.* when they are "at rest"), discharge synchronously about 10 times per second; their pattern of "inactivity," so to say, is a mark-time beat. Purposive activity of the cerebral cortex is associated with a *complicated* pattern of discharge of the cells involved, which gives rise to an undecipherable record when the usual indirect leads are employed.

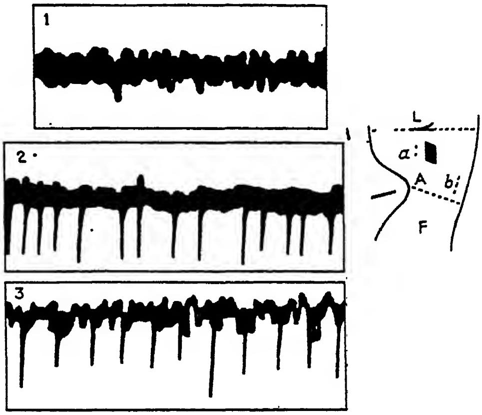


FIG. 390.—Physiological Neuronography. (Dusser de Barenne and McCulloch, *J. Neurophysiol.*, 1938, *1*, 73.)

The inset shows the motor area (area 4) of the cerebral cortex of *Macaca mulatta*. L, A, F represent leg, arm, and face areas respectively. The small black square is the area to which the strychnine solution was applied. The small dots, labelled *a* and *b* respectively, are the regions in the arm area (A) from which the electrical potentials were recorded.

1. Control electrocorticogram from arm area. 2, 3. Electrocorticogram from region *a* and region *b* respectively, at the height of "strychnine spikes" (the spikes are the long heavy down strokes in the record). These records prove that the axons from the stimulated area re-enter the cortex at *a* and *b*; in the technical terminology employed, areas *a* and *b* have been "fired" from the stimulated region.

The clinical significance of electroencephalographic records is considered in detail on p. 624. It is first necessary to discuss the fundamental question of how cortical potentials are produced under the resting conditions described above.

Mode of Production of "Resting" Cortical Potentials.—Two possible explanations will be considered: (*a*) the cortical cells possess an inherent rhythmicity (*e.g.* like the sino-auricular node) and, when not subjected to external influences, spontaneously generate impulses at a regular rate; (*b*) the discharge of the so-called "resting" cortical cells is due to a regularly recurring stimulus coming from another part of the brain. Experimental evidence shows that the "resting" cortical discharge is due to the existence of a *closed circuit consisting of cortex→thalamus→cortex*; the cortex stimulates the thalamus which in turn stimulates the cortex again; the cortical cells discharge every time the circuit is traversed. Hitherto psysiologists have

attempted to explain all types of neural activity solely in terms of reflex arcs of varying degrees of complexity. The closed circuit, which may also be of varying complexity, represents a conception of neural organization, hitherto

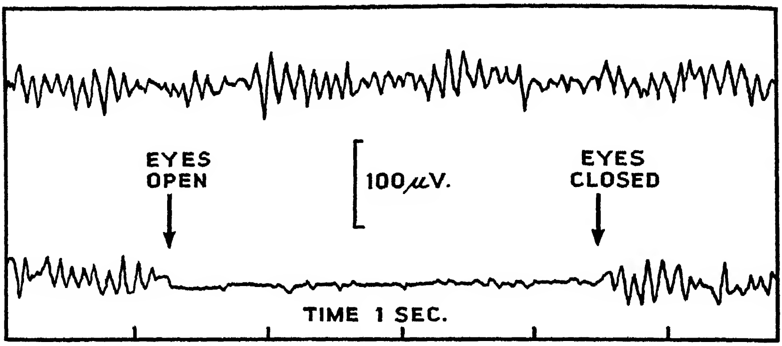


Fig. 391.—Normal Human Electroencephalogram (EEG).

A. Normal record taken with eyes closed, showing alpha waves.
B. Normal record. On opening the eyes the regular alpha rhythm is replaced by small irregular oscillations. The normal waves return on closing the eyes. Time in seconds.

unrecognized, and one of great interest and importance. The evidence is as follows:

(i) An *electrocorticogram* is obtained in animals by leading directly from the exposed surface of the cerebral cortex; the record (Fig. 392) resembles in a general way the indirectly obtained electroencephalogram, in that it consists of a more or less regular series of potential waves. In the motor area these waves are mainly generated by layer V; thus if the outer four layers of the motor cortex are destroyed, the record remains substantially unaltered in character; if layer V is also destroyed, the frequency of the waves becomes lower and their voltage becomes smaller.

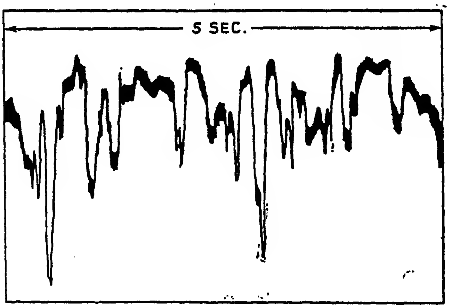


Fig. 392.—Electrocorticogram from Motor Arm Area (A4) of Brain of Macaque. (Dusser de Barenne and McCulloch, *J. Neurophysiol.*)

Note that this record is taken directly from the surface of the cortex.

(ii) (*a*) The potential waves in a restricted area are unaffected by a *circular cut* which isolates the region of cortex under examination from the rest of the cerebral grey matter; on the other hand, *undercutting* the cortical area, that is, severing the white fibres entering and leaving it, abolishes the waves, proving that they are not generated "spontaneously."

(*b*) Transection of the brain stem at the level of the medulla or midbrain does not significantly modify the waves; but they are abolished by a deep cut through the cerebral white matter ***severing the connections between the cortex***

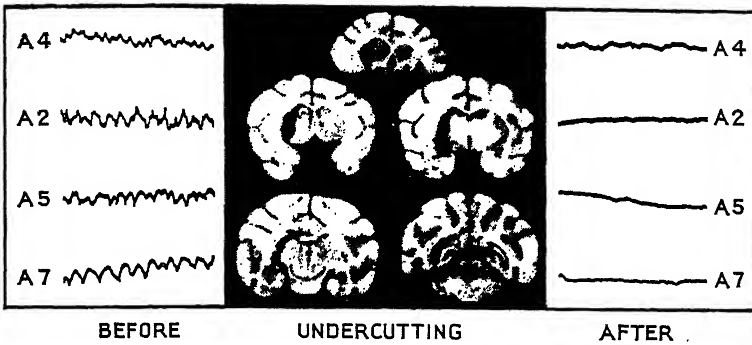


FIG. 393.—Severing Thalamo-cortical Connections Abolishes the Resting Electrocorticogram. (Barenne and McCulloch, *J. Neurophysiol.*, 1941, *4*, 307.)

Experiment on macaque under Dial anæsthesia.
Records on left and right are taken respectively before and after deep undercutting of the sensori-motor cortex (as illustrated in the central panel), severing all thalamo-cortical connections on that side. A4, A2, A5, A7 are records taken from arm areas 4 (motor), 2, 5, 7 (sensory). Note the normal electrical activity in the control records. After the undercutting operation, all electrical activity is abolished.

and the thalamus (Fig. 393). The resting electrocorticogram thus depends on the integrity of the connections between cortex and thalamus.

(*c*) A further clue is provided by the following experiment. In certain states of anæsthesia, the cortical potential waves occur in groups separated by intervals of inactivity. Appropriate stimulation of the *medial* thalamus during these silent intervals can set up typical runs of alpha waves in the cortex (Fig. 394).

The experimental results demonstrate the existence of a closed circuit, *i.e.*

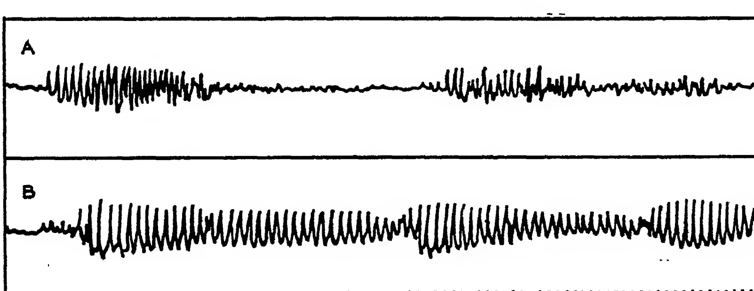


FIG. 394.—Stimulation of Medial Thalamus Produces Cortical Potentials Resembling the Normal Resting Electroencephalogram or Electrocorticogram. (Morison and Dempsey, *Amer. J. Physiol.*)

Lightly anæsthetized cat: electrocorticogram from middle suprasylvian gyrus.
A. Spontaneous electrical waves at frequency of 8–12 per second, appearing in bursts and alternating with periods of feeble electrical activity.
B. During the period indicated by the signal (lowest line) the medial thalamus was stimulated at the frequency shown by the signal. Electrical waves develop in the cortex, resembling those of the normal electroencephalogram; the waves wax and wane in magnitude, and show a uniform frequency. These waves appear all over the cerebral cortex; they appear in areas of cortex isolated except for their thalamo-cortical connections.

thus, one which receives impulses from the strychnine-stimulated region). But if leads are taken from any other cortical area it is found that stimulation of the suppressor band causes no impulses to be sent to other cortical areas (*i.e.* these areas are not fired); but, as mentioned, the "resting" cortical potentials in these regions are extinguished (disappearance of the cortical potentials is shown in the "firing diagram" by a horizontal line, thus —) (Fig. 396). Fig. 397 for example shows that stimulation of the suppressor band 4s depresses the electrical activity of the motor cortex.

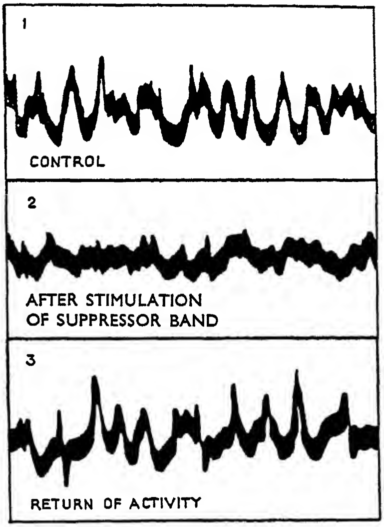


Fig. 397.—Suppression of Electrical Activity in Motor Cortex as a Result of Stimulation of Suppressor Band. (Barenne, McCulloch, and Ogawa, *J. Neurophysiol.*, 1938, *1*, 438.)

Experiment in macaque under Dial anæsthesia. The records are taken from the face area of the motor cortex (area F4).

1. Control record showing resting activity.
2. Record taken 11 minutes after application of strychnine to the suppressor band 4s. Note the great reduction in the electrical activity; the tall waves have disappeared.
3. Record taken 3 minutes later, showing return of electrical activity to the control level.

Mechanism of Suppression.—The mechanism is the same for all the suppressor bands; impulses pass down from the suppressor band to the *caudate nucleus* (Fig. 398) which is "fired." From the caudate nucleus, impulses pass to the thalamus which is inhibited, as shown by the fact that the resting thalamic potentials are depressed. Inhibition of the thalamus blocks the passage through it of the normal cortico-thalamo-cortical impulses which maintain the resting cortical potentials. These cortical potentials consequently disappear (Figs. 399, 408).

It seems probable that all regions of the cortex maintain closed excitatory circuits with the *medial* nuclei of the thalamus. The *lateral* nucleus of the thalamus, on the other hand, merely relays impulses from the skin and muscles of the face, arm, and leg to the postcentral cortex (Fig. 357). The suppressor path from the cortex

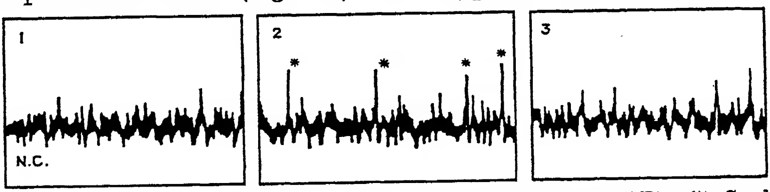


Fig. 398.—Stimulation of Suppressor Band Sets up Activity in ("Fires") Caudate Nucleus. (Barenne and McCulloch, *J. Neurophysiol.*, 1938, *1*, 367.)

Records are taken from an electrode in the caudate nucleus.

1. Control record.
2. Record taken 5 minutes after applying strychnine to the suppressor band A4s (*i.e.* the arm region of this band). The large upward spikes indicated by asterisks represent activity developing in the caudate nucleus ("firing" of caudate nucleus).
3. Record taken 10 minutes later; the induced activity has subsided.

via the caudate nucleus to the thalamus normally maintains a *continuous* inhibitory influence on the level of the "resting" cortical activity. Thus an acute destructive lesion of the caudate nucleus (severing the

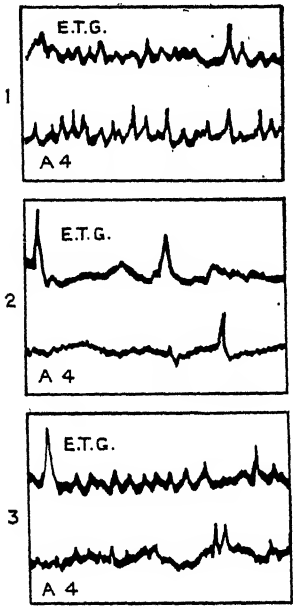


FIG. 399.—Stimulation of Caudate Nucleus Suppresses Electrical Activity in the Thalamus and in the Cerebral Cortex. (Barenne and McCulloch, *J. Neurophysiol.*, 1938, *1*, 372.)

Experiment on macaque under Dial anæsthesia.
In 1, 2, and 3 the records from above downwards are: E.T.G.: electrical activity of nucleus of thalamus (electrothalamogram). A4: electrical activity of arm area of motor area 4.
1: control records showing resting activity.
2: records taken after stimulation of caudate nucleus by local application of strychnine. Note that there are long periods in the records in which electrical activity is almost completely suppressed
3: record taken later in which the electrical activity has returned to its control level.

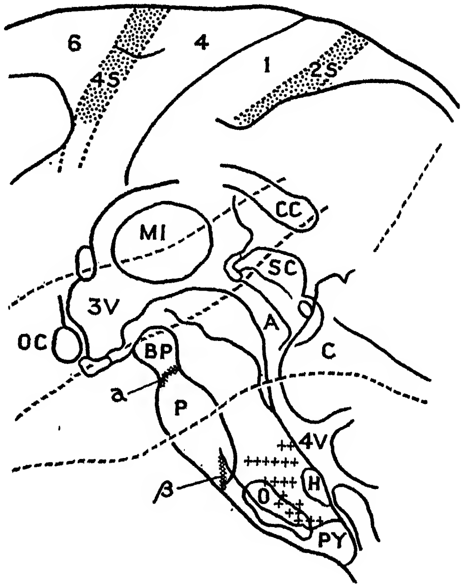


FIG. 400.—Descending Inhibitory Pathway from Suppressor Bands via Reticular Nuclei of Brain Stem to Spinal Cord. (After McCulloch *et al.*, *J. Neurophysiol.*, 1946, *9*, 130.)

Midsagittal reconstruction of brain stem (monkey) with cerebral cortex superimposed (the parts of the cerebral cortex which are treated as transparent are indicated by dotted lines).
4, 6, precentral motor cortex. 1, postcentral sensory cortex. 4s, 2s, cortical supressor bands. MI, massa intermedia. CC, corpus callosum. 3V, third ventricle. OC, optic chiasma. SC, superior colliculi. BP, basis pedunculi. P, pons. A, aqueductus Sylvii. C, cerebellum. 4V, fourth ventricle. O, olive. H, hypoglossal nucleus. Py, pyramid. +++, the region of brain stem (*nuclei of reticular formation*) which is "fired" (activated) when the cortical suppressor bands 4s, 2s are stimulated. *Stimulation of the brain stem region* +++ (*reticular nuclei*) *inhibits the spinal motor neurones.* The descending fibres from 4s and 2s run in the pyramidal tract in the basis pedunculi and pons; they are cut by a section through the basis pedunculi at α. They end in the reticular formation of the pons and medulla; the next relay, the reticulospinal fibres, are *not* cut by a section at β through the upper medullary pyramid.

suppressor pathway) is followed by the development of *exaggerated* electrical activity both in the thalamus and in the cerebral cortex. It follows that impulses from any source that modify the excitability of the medial thalamic nuclei may alter the amplitude, frequency, and other

characters of the resting electrocorticogram or electroencephalogram. Stimulation of the hypothalamus (surprisingly) sends impulses up into the thalamus and alters the pattern of cortical activity. The waves in the electroencephalogram may thus be expected to show characteristic features in different individuals according to their emotional or mental make-up, or the state of health (or disease) of the organs in general.

(2) DESCENDING INHIBITORY PATHWAY TO LOWER MOTOR NEURONES. — From the suppressor bands descending fibres arise which give collaterals to the caudate nucleus and receive descending fibres from that nucleus. The fibres from the cortex and caudate nucleus end in the *nuclei of the reticular formation* throughout the brain stem (p. 702; Fig. 400). From these nuclei, reticulospinal fibres arise which end round the cranial and spinal lower motor neurones and *inhibit their activity*.

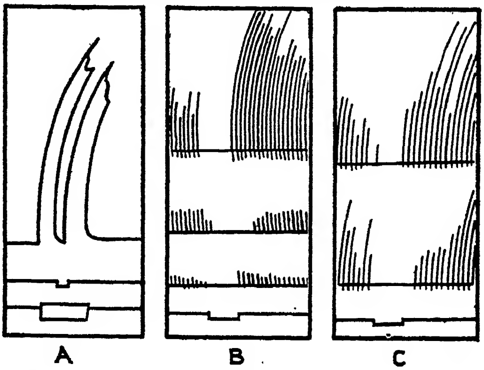


FIG. 401.—Descending Inhibitory Pathways from Nuclei of Reticular Formation of Brain Stem. (Magoun and Rhines, *J. Neurophysiol.*, 1946, *9*, 165.)

A. During lower signal: stimulate internal capsule (pyramidal tract) to produce flexion of hind limb. During upper signal: stimulate reticular nuclei in brain stem. Temporary inhibition of muscular contraction.
B. Records from above downwards: knee-jerk; flexor reflex; blink reflex (elicited at 2 sec. intervals); signal line. During signal: stimulate reticular nuclei. All the reflexes are inhibited. In the case of the knee-jerk there is post-inhibitory rebound.
C. Records from above downwards: contractions of fore limb and contraction of hind limb elicited by stimuli applied at short intervals to motor cortex; signal line. During signal: stimulate reticular nuclei. Temporary inhibition of responses followed by some rebound.

The functions of this inhibitory pathway are illustrated by the following observations.

(i) Stimulation of the suppressor area 4s abolishes muscle tone and inhibits any movements which are taking place at the time. Stimulation of the relay nuclei in the reticular formation likewise produces inhibitory effects; it inhibits reflexes involving the face (*e.g.* blink reflex), the arm or the leg (*e.g.* knee-jerk) (Fig. 401).

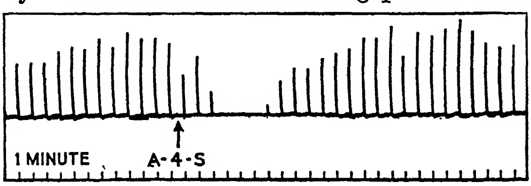


FIG. 402.—Stimulation of Suppressor Band Annuls Effect of Stimulating Motor Cortex. (Barenne and McCulloch, *J. Neurophysiol.*, 1941, *4*, 315.)

Stimulate arm area (A4) of motor cortex to produce wrist extension at intervals of one minute throughout the Fig. At the arrow, *brief* stimulation of suppressor area 4s. Note that the response to motor cortex stimulation rapidly disappears and then slowly recovers. Time in minutes.

(ii) A cortical motor point is stimulated at regular intervals to produce a movement, *e.g.* wrist extension. While this background activity is taking place the suppressor area 4s is stimulated for a short time. After a latent period, the muscular contractions in response to motor cortex stimulation,

diminish and ultimately disappear; recovery of the response gradually develops (Fig. 402). It may be concluded that the inhibitory pathway from the suppressor area to the spinal cord competes with and gradually annuls the excitatory effect at the spinal motor neurones produced by impulses passing in the pyramidal pathway.

(iii) Extirpation of area 4s releases the lower motor neurones from a tonic inhibitory influence thus giving rise to increased muscle tone of characteristic distribution (spasticity) and exaggerated deep reflexes. These changes are similar to those observed in clinical hemiplegia. The inhibitory fibres from area 4s are probably intermingled with the pyramidal tract fibres in the internal capsule, in the midbrain [basis pedunculi] and in the pons. The next relay, the inhibitory reticulospinal fibres, are *not* associated with the pyramidal fibres in the medulla (Fig. 400), but may rejoin them in the spinal cord. It is likely that the changes in posture and reflexes in clinical hemiplegia are not due to injury to the pyramidal tracts but to involvement of the inhibitory pathway (p. 645).

As area 4s is so closely associated anatomically with the motor area (Fig. 407) the results of stimulation or extirpation of the motor area may be seriously modified by accidental involvement of the suppressor area.

CLINICAL ELECTROENCEPHALOGRAPHY [1]

The objects of clinical EEG recording are :—

(i) to obtain an idea of the distribution of the electrical activity over as wide an area of the cortical surface as possible;

(ii) to observe, simultaneously, activity arising in different parts of the brain;

(iii) to obtain a continuous recording for as long a period as possible from various parts of the brain so as not to miss abnormal activity which may occur only sporadically;

(iv) to observe the response to certain standard test situations (*evoked activity*). To attain these objects it is desirable to use a multiple amplifier EEG machine which enables several records to be obtained simultaneously. Usually six simultaneous tracings are taken. Many electrodes are placed on the scalp and connected in pairs to the EEG machine so as to give antero-posterior and transverse " cross sections "of the brain in all areas. Each tracing (Fig. 403) is a complex rhythmic wave which seldom, if ever, repeats itself precisely.

It is an obvious preliminary to the classification of abnormal rhythms that there should be some method of analysing and classifying the " normal " record in all its range of variation. This difficult undertaking can be accomplished instrumentally. The analyser most widely used in Britain makes, automatically, an analysis of the EEG tracing every 10 seconds, and writes out a histogram of the wave components underneath the complex wave tracing to which the analysis refers (Fig. 403). The analyser gives the answer in terms of the relative amplitudes of the pure wave (sine wave) components present in the complex wave analysed. This form of analysis is known as

[1] Res. Publ. Assoc. nerv. ment. Dis., *Epilepsy*, 1947. Gibbs and Gibbs, *Atlas of Electroencephalography*, 1941; vol. I, 2nd edn., Massachusetts, 1950. Hill and Parr, *Electroencephalography*, London, 1950.

Section contributed by W. F. Floyd.

Fourier analysis and is applicable to any complex recurring waveform (*e.g* sound waves; the cochlea is an acoustic analyser, operating in a similar manner).

For convenience of description the frequency spectrum from 1 to 30 cycles per sec. (c/s) is divided into five bands as shown in the following table. This range covers all the frequencies present in most EEG records to which attention has so far been paid.

Frequency c/s.	*Band.*
1–3·5	delta (δ)
4–7	theta (θ)
8–13	alpha (α)
14–18	beta (β)
20–30	gamma (γ)

The alpha rhythm is the main rhythm ordinarily found in the normal adult; its amplitude is some 50 μV. It is usually a complex wave with two

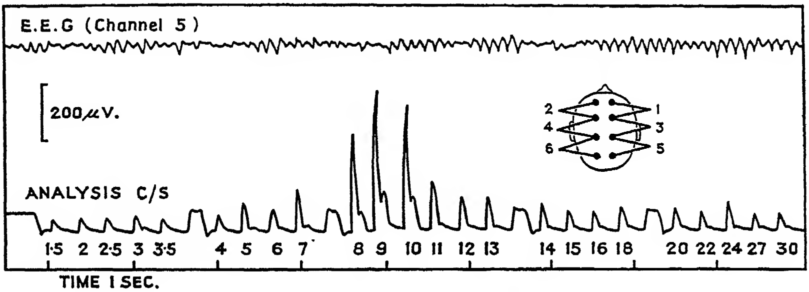


FIG. 403.—Normal Human Electroencephalogram and its Analysis.

Upper record: E.E.G., Channel 5.
The disposition of the surface electrodes on the head is shown in insect (top = front, bottom = back of head).
Lower record: Analysis of Channel 5.
The height of each wave is proportional to the magnitude of the contribution of the particular wave frequency.

or three pure wave components, all of which lie within the alpha band (as defined above), but the relative amplitudes of the components may vary continuously. The alpha rhythm is present maximally in the occipital and parieto-occipital area, and is usually diminished during visual activity and mental effort. The alpha rhythm as thus defined is not simply rhythmic activity falling in a particular frequency range (8–13 c/s), but it involves cortical localisation and a particular response characteristic. However, the term "alpha rhythm" is sometimes used more loosely to describe any rhythmic activity the frequency of which lies within the alpha range.

In some adults a low amplitude (10 μV) theta rhythm is also found, usually over the parietal and temporal areas. During sleep, delta waves may appear, and may reach an amplitude of 100 μV, and are found diffusely over all areas of the brain.

In children, electrical activity is present in the brain from birth onwards, but can only be recorded at first during sleep. From 3 months of age onwards, rhythmic activity is detectable in the occipital region, at first at 3–4 c/s and increasing in frequency as the age advances. The normal adult rhythm is

established by 8–10 years. In old age the maturation process is reversed, senility being accompanied by some slowing of the rhythm.

Evoked Activity.—Three methods are commonly used in clinical work to evoke EEG response patterns.

(i) *Overventilation.* The response of the normal subject to overventila-

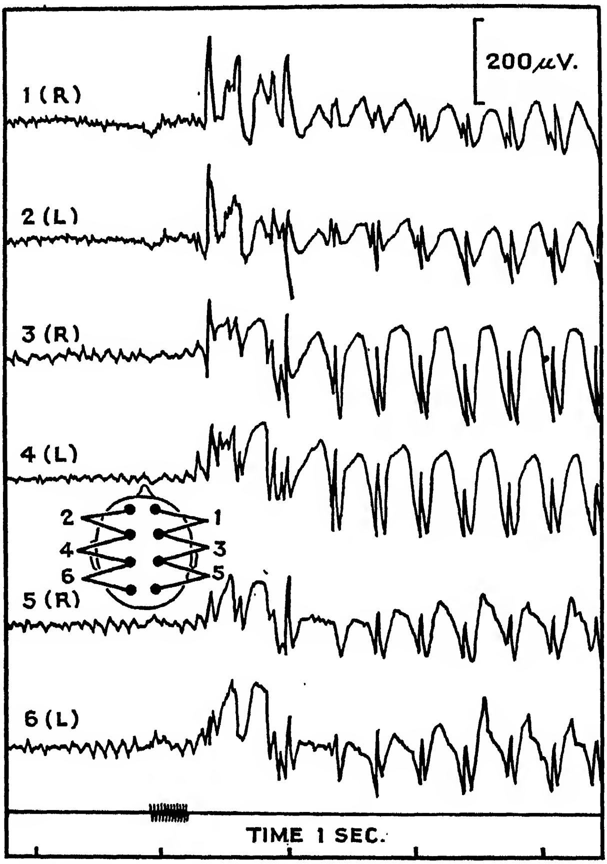


FIG. 404.—Electroencephalogram in an attack of *Petit Mal.*

Records from above downwards: channel 1 (right), 2 (left), 3 (right), 4 (left), 5 (right), 6 (left). Inset shows disposition of surface electrodes on head (1, 2, anterior; 5, 6, posterior).

Left-hand part of records: normal waves. At signal: attack sets in. Note " Spike and Wave " in all leads.

tion for several minutes is often nil, but wide variation is seen. The procedure may evoke high voltage delta activity of 2–3 c/s: sometimes theta activity (4–7 c/s) is also present. The delta activity appears in 1–2 minutes and ceases within a few seconds of cessation of overventilation. The clinical test is usually a 3–4 minute period of overbreathing (marked increase of depth but not usually of rate).

(ii) *Photic Stimulation.* A powerful flashing light is directed into the eyes

and the frequency of flashing is varied. Normal subjects show a variety of responses, of which one of the commoner is the appearance of the flicker frequency as a component of the EEG rhythm.

(iii) *Chemical Stimulation.* Certain drugs, *e.g.* leptazol (pentamethylenetetrazole), or insulin, are administered, usually intravenously, in an attempt to diagnose epilepsy by the abnormal rhythms so evoked in epileptic subjects.

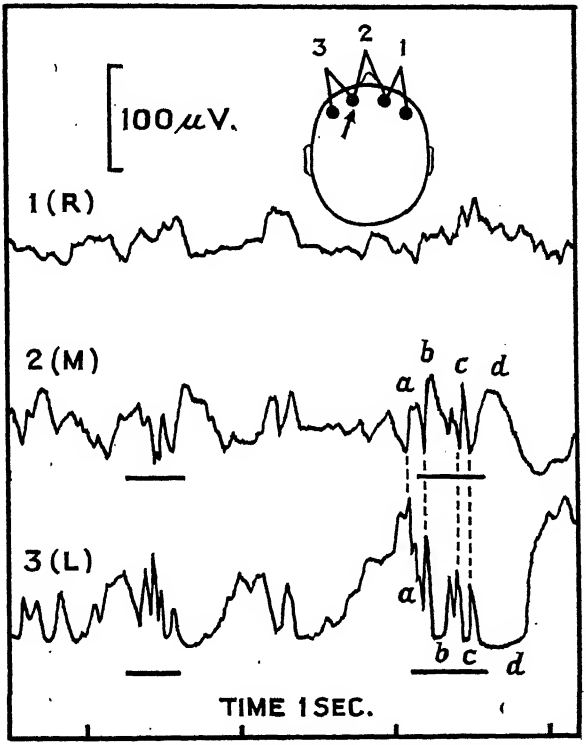


Fig. 405.—Electroencephalogram showing "Focus."

Inset: Disposition of electrodes across anterior part of head.
Records from above downwards: Channel 1 (right), 2 (midline), 3 (left). Examine part of channels 2 and 3 marked by the right horizontal line. Corresponding waves have been indicated by the connecting dotted lines. Note that each wave in channel 2 is similar in shape to the wave in channel 3, but is *opposite in direction.* This finding demonstrates that the abnormal focus is under the electrode which is common to channels 3 and 2 (marked by arrow in inset).

Abnormal Rhythms.—Cerebral dysrhythmia is found in a variety of pathological states, of which the more important for diagnostic purposes are: the epilepsies, local cerebral lesions, (*e.g.* abscess, tumour, or trauma), infective conditions and degenerative processes. The spike and wave complex ("dart and dome" of the American literature) is the commonest feature of epileptic records (Fig. 404). The simplest type of spike and wave complex consists of a single spike associated with each wave; the waves

occur most often at about 3 per sec. The spike amplitude is variable over different areas of the brain, but the wave amplitude is more constant and may reach ·1 mV or more. The complex occurs spontaneously in the resting record of epileptic subjects (most often in *petit mal*). As a single isolated complex appearing in one or more leads it is usually unaccompanied by any change in consciousness: as a repeated pattern of cerebral discharge it is generally accompanied by a minor attack with the usual clinical manifestations, such as, fixation of head (*e.g.* rotation to left or right), eyes open and staring, and transient loss of consciousness. The repeated spike and wave pattern appears also as a response to overventilation in some epileptics. Its presence in an EEG record is diagnostic of epilepsy, but not necessarily of a particular clinical entity such as *petit mal*: other features of the record and the patient's clinical history have to be taken into account in making a diagnosis. It is also found in some patients with cerebral injury or tumours, but usually in these cases the complex is strictly localised to the vicinity of the lesion.

In about 50% of patients subject to major seizures, the inter-seizure EEG

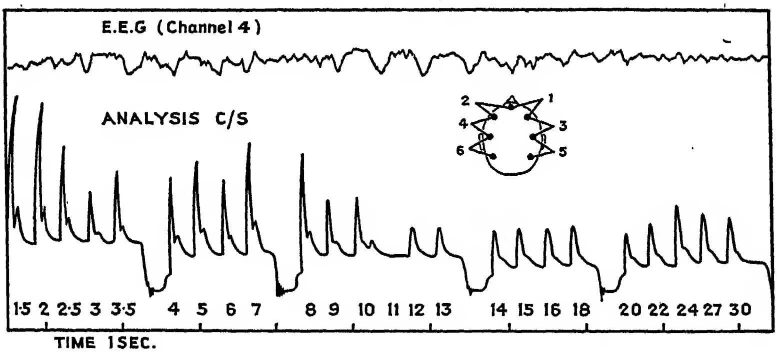


FIG. 406.—Electroencephalogram from Case of Intracranial Tumour, showing Preponderance of Slow Wave Activity.

Upper Record: EEG, channel 4. Lower Record: analysis of channel 4. Compare normal findings in Fig. 403).

record is normal. In the remainder the abnormal features, including the overventilation record, cover a wide variety of patterns, *e.g.* the presence of delta waves, theta waves, or fast activity (mainly beta).

It is rare for an EEG recording to coincide with a major fit. Such records of the major fit as have been made contain high frequency components (mainly in the form of "spikes") during the tonic and clonic stages of the seizure, preceded and followed by a rhythm containing both delta and theta activity components.

EEG records containing typical epileptic wave patterns (*e.g.* spike and wave complexes or paroxysmal spike discharges) are found in certain cases of cerebral tumour or trauma. This epileptiform activity is usually restricted to the area of the brain involved in the lesion and the abnormal activity shows a "focus" area, *i.e.* the wave pattern appears on each side of the focus but in opposite polarity (Fig. 405), thus producing a reversal of the phase of the spikes or the waves over the site of origin ("focus").

The typical EEG record of an intracranial tumour is one in which slow wave activity is the dominant feature; the record is often of high voltage

(100 μV or more) and the delta rhythm component shows a focus in the vicinity of the tumour. In addition there may be a theta rhythm present (but usually only if the tumour is deep seated) as well as the normal alpha rhythm components. The delta waves are often quite irregular in form and occurrence, but may be present as a semi-continuous train of constant amplitude almost sinusoidal waves (Fig. 406). The source of the abnormal delta rhythm is not the lesion itself but the adjoining cerebral tissue which is under the mechanical stress set up by the tumour tissue. The effect of mechanical pressure on normal cortical tissue is to evoke slow wave activity. The greater the rate of increase of pressure, the slower the induced rhythm and the greater its amplitude. Small, slowly growing tumours (*e.g.* benign tumours) are not so readily detected, therefore, by electroencephalography as are rapidly growing tumours (*e.g.* an astrocytoma).

EXCITOMOTOR AREAS. PYRAMIDAL TRACTS. CLINICAL HEMIPLEGIA [1]

Excitomotor Areas.—The term excitomotor area is generally applied to that part of the cortex of the frontal lobes which on stimulation gives rise to skeletal muscle responses. As this region lies anterior to the central sulcus it is commonly called the *precentral motor cortex.* It is divisible into several cytoarchitectonically distinct zones (Fig. 407) with fairly distinctive functions.

(1) Area 4.—This occupies almost the whole length of the precentral gyrus; it extends back into the central sulcus itself and over on to the medial surface of the hemisphere.

Area 4 is the main region of origin of the *pyramidal tracts* (some of these fibres, however, also arise from area 6). Stimulation of this region produces co-ordinated movements of the opposite face, arm, and leg; area 4 is thus frequently called the "true" motor cortex (p. 633). The pyramidal tracts are fully considered on p. 632.

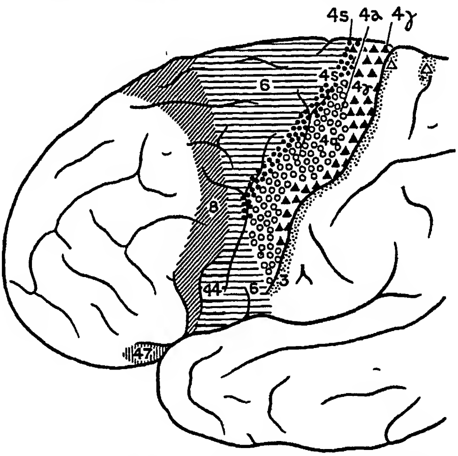


Fig. 407.—Diagram of Precentral Motor Cortex in Man. (After Bonin, in Bucy, *Precentral Motor Cortex*, University of Illinois Press, 1944.)
Area 4 is divided into three zones: 4γ, contains "giant" Betz cells. 4α, contains no giant Betz cells. 47=area 13 (Walker). △=Betz cells in postcentral area. 4s=suppressor band. 8, contains frontal eyefield and suppressor band 8s.

(2) Area 6.—This area mainly gives rise to a long descending relaying *excitatory* pathway sometimes called the *extrapyramidal* tract; (some of these fibres also arise in area 4). The fibres relay successively in the following

[1] Res. Publ. Assoc. Res. nerv. ment. Dis., *Localization of Function of Cerebral Cortex*, 1934; *Frontal Lobes*, 1948. Hines, *Biol. Rev.*, 1943, *18*, 1. Bucy, *Precentral Motor Cortex*, 1944; 2nd edn., 1949. Walshe, *Brain*, 1942, *65*, 409.

grey masses (Fig. 408) putamen and globus pallidus (of lenticular nucleus); thalamus; hypothalamus; nuclei of reticular formation in the dorsal part of the brain stem; the last relay, the reticulospinal fibres, ends round the lower motor neurones. The following experiments illustrate the probable function of these fibres.

(i) After destruction of the pyramidal tract, stimulation of area 6 is said to produce rotation movements of the head, eyes, and trunk in the opposite direction, and the opposite limbs carry out complex co-ordinated movements

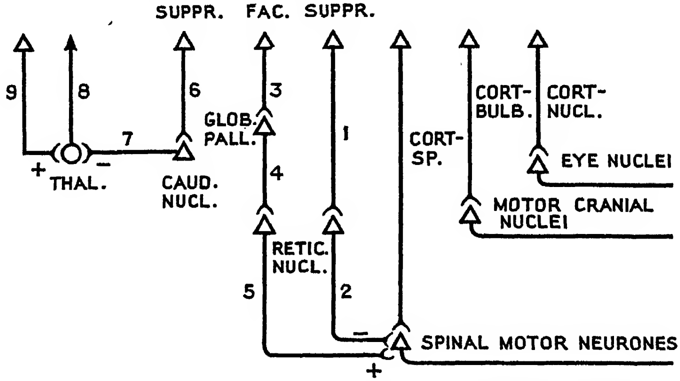


FIG. 408.—Descending Pathways from Precentral Motor Cortex.

Cort.Nucl. = Corticonuclear fibres (from frontal eyefield).
Cort.Bulb. = corticobulbar fibres (from "face" area).
Cort.Sp. = corticospinal fibres (pyramidal tracts).
Suppr. = suppressor bands (*e.g.*, 4s, 8s).
1, "inhibitory" pathway from suppressor band to reticular formation stimulating nuclei there.
2, reticulospinal fibres *inhibiting* spinal motor neurones. Section of 1 or 2 produces "spasticity."
Fac. = cortical facilitatory area (in area 6).
3, 4, 5, descending facilitatory pathway stimulating spinal motor neurones.
3, area 6 to globus pallidus, etc.
4, globus pallidus, thalamus and hypothalamus, to reticular formation.
5, reticulospinal fibres, which *excite* the spinal motor neurones.
Thal = thalamus.
Caud.Nucl. = caudate nucleus.
9, fibres from cortex to thalamus.
8, thalamo-cortical fibres.
9, 8, this circuit maintains resting electrocorticogram.
6, from suppressor band to stimulate caudate nucleus.
7, from caudate nucleus to inhibit thalamus.
6, 7, suppresses resting electrocorticogram.

of a flexor or extensor character. These responses may be mediated by this extra-pyramidal path.

(ii) Stimulation of the hypothalamus facilitates the response to motor cortex stimulation. The experimental procedure is as follows: the excitomotor cortex is stimulated weakly to produce, say, slight flexion of the ankle of one hind limb. The cortical stimulus is then *combined* with appropriate hypothalamic stimulation (the latter alone is without effect on movement); the *combined* stimulation gives rise to movements which are more powerful and widespread, *e.g.* the response may now consist of powerful flexion of all the joints of the hind limb accompanied by flexion of the fore limb (cf. Fig. 409, A).

Hypothalamic stimulation similarly facilitates the response to stimulation

of the medullary pyramid even *after removal of the entire cerebral cortex.* The hypothalamus thus produces its effect by means of a *descending* pathway to the lower motor neurones.

(iii) Facilitation of the effects of pyramidal tract stimulation is also obtained on stimulating the thalamus and the reticular formation of the brain stem, in other words, any part of the facilitatory pathway detailed above.

(iv) Hypothalamic stimulation may facilitate the knee-jerk (Fig. 409, B).

(v) This facilitatory mechanism is possibly damaged in Parkinson's disease; the excitability of the lower motor neurones would then be decreased leading to the characteristic weakness and paucity of voluntary movement (p. 659).

(3) Area 8.—The part of area 8 which is mainly in the middle frontal convolution is the *frontal* eyefield which controls eyeball movements through its descending *corticonuclear tract* (Fig. 408). These fibres descend in the anterior limb of the internal capsule, then in the medial fifth of the pes pedunculi and finally pass dorsally to supply the eye nuclei of the opposite side (p. 637).[1]

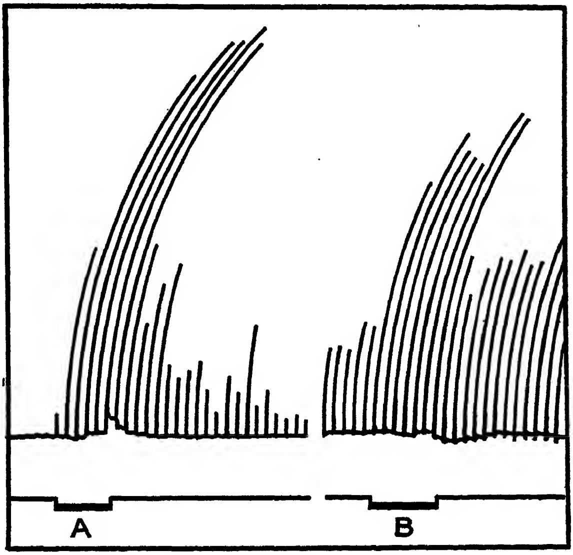


Fig. 409.—Descending Facilitatory Pathways from Hypothalamus and Brain Stem to Spinal Cord. (Rhines and Magoun, *J. Neurophysiol.*, 1946, *9*, 219.)

Monkey.
A. Throughout the record, stimulate motor cortex at 2 sec. intervals; initially there is no response of the limb.
During descent of signal, A, stimulate hypothalamus; cortical impulses are facilitated (at the ventral horn cells) and the limb contracts. (Stimulate hypothalamus alone: no effect on limb.)
B. Throughout the record, elicit knee-jerk at 2 sec. intervals.
During descent of signal, B, stimulate hypothalamus The knee-jerk is facilitated (the response is enhanced).

(4) Areas 8s and 4s (Fig. 395).—These are two of the suppressor bands (p. 620).

Two other important connections of the excitomotor cortex should be mentioned.

(1) *Cortico-Pontine Fibres.*—These arise mainly in areas 4 and 6 (but also from the temporal, occipital, and parietal lobes) to end round the nuclei pontis; the impulses are relayed via the brachium pontis to the neocerebellum of the opposite side.

(2) *Cortico-Thalamic Fibres.*—As from most other parts of the cortex,

[1] Another area controlling eyeball movements is found in the occipital lobe (areas 17 and 18) (*occipital eyefield*).

fibres descend from the excitomotor area to the thalamus forming part of the closed circuit consisting of cortex–thalamus–cortex.

Some of the principal descending fibres from the precentral cortex are shown diagrammatically in Fig. 408.

Pyramidal Tracts.—These pass from the motor area to the spinal ventral horn cells and to all the motor cranial nuclei except those supplying the external eye muscles. The pyramidal tracts and their cells of origin constitute the *upper motor neurone*; the spinal and cranial motor neurones constitute the *lower motor neurone* (p. 499). The pyramidal tract fibres to the spinal ventral horn cells constitute the *corticospinal tract*; the fibres to the motor cranial nuclei constitute the *corticobulbar tract.* In man the pyramidal tract fibres number about one million, of which 40% are non-myelinated. The diameter of the myelinated fibres varies between 1 and 20 μ; the great majority (over 80%) are 3 μ or less, only 4% are over 10 μ and 2% are 11–20 μ (Fig. 410). As the velocity of conduction in nerve fibres is directly proportional to their diameter (p. 492) it follows that the pyramidal tract is overwhelmingly a slowly conducting pathway. The pyramidal tracts arise from pyramidal shaped cells of all sizes (both large and small)[1] situated in layer V of the motor area, *i.e.* area 4; but some of the fibres however arise in area 6 (Fig. 407).[2]

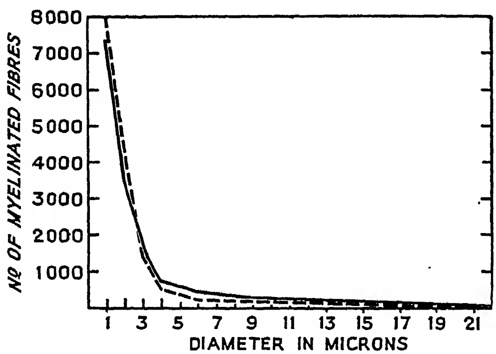


FIG. 410.—Distribution of Diameter of Myelinated Fibres of Corticospinal Tract in Man. (Lassek, *J. comp. Neurol.*, 1942, *76*, 219.)

COURSE OF THE PYRAMIDAL TRACTS (Figs. 411, 412).—The pyramidal tracts converge from the precentral cortex through the corona radiata to reach the *internal capsule.* This is a mass of white fibres lying between the basal ganglia, limited laterally by the lenticular nucleus and medially by the caudate nucleus and thalamus. In horizontal section the internal capsule is V-shaped, the point of the V looking medially. The pyramidal tracts lie in the bend [the genu], and the anterior two-thirds of the posterior limb [occipital part]. The fibres from before backwards are concerned with the control of head, shoulder, elbow, wrist, fingers, trunk, hip, knee, and toe movements in the order named. (It should be noted that *extrapyramidal* fibres, both excitatory and inhibitory are *intermingled* here with the pyramidal fibres). Immediately behind the pyramidal tracts lies the condensed *sensory* path, and the *visual* path a little farther back. More posteriorly still lie the *auditory* fibres and the temporopontine tract. In the anterior limb [frontal part] are found fibres from the frontal lobes mainly (and also from the

[1] The largest pyramidal cells in layer V of area 4 are called Betz cells; they are most numerous in the leg area and fewest in the face area. They only number 35,000 in all and thus can only give rise to 5% of the fibres of the corticospinal tract.

[2] Some workers claim that the cortical area of origin is wider and that some of the pyramidal fibres arise from the parietal and temporal lobes also.

parietal and temporal lobes), to the basal ganglia, the eye nuclei (cortico-nuclear fibres, p. 637), and other regions of the brain stem (especially the pons). Obviously a small injury in this region can produce most widespread motor and sensory disturbances (Fig. 411).

In the *crus* the pyramidal fibres lie ventral to the substantia nigra, occupying the middle three-fifths of this region. [The medial fifth carries the cortico-nuclear and frontopontine, and the lateral fifth the temporopontine, fibres.] In the *pons* the pyramidal fibres are broken up into a series of scattered bundles by the nuclei pontis and the crossing fibres of the brachium pontis. Stimulation of the cut surface of the crus and pons in the ape shows that there is well-marked localization of the pyramidal fibres for different parts of the body (Fig. 412, I). From without inwards the order of the fibres is foot, hip and knees, abdomen and chest, fingers and wrist, face and tongue. Throughout the brain stem, the *corticobulbar* fibres are crossing to reach the motor cranial nuclei of the opposite side. In the *medulla* the corticospinal fibres reunite to form a compact ventrally projecting mass, the *pyramid.* The pyramidal tracts were so named because they were recognized as the *tract in the pyramid* (and *not* because they arise in pyramidal shaped cells in the cerebral cortex). In the lower part of the medulla, the main pyramidal decussation takes place (Fig. 412, III). Most of the fibres cross over to the opposite side and dorsally, to come to lie in the lateral columns of the spinal cord as the crossed pyramidal [lateral corticospinal] tract. Some (about 15% of the fibres) stay on their own side and in their original position and continue into the cord close to the anterior median fissure as the direct pyramidal [anterior corticospinal] tract (Fig. 412, IV). The crossed fibres ultimately connect directly and via dorsal interneurones with ventral horn cells.

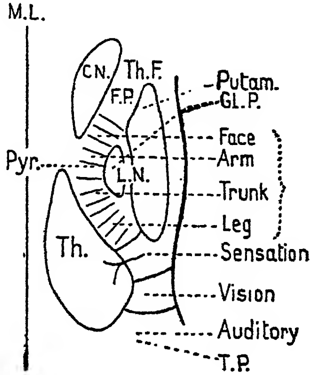


FIG. 411. — Horizontal Section Through the Base of the Brain to Show the Basal Ganglia and the Internal Capsule (Diagrammatic).

M.L.=Middle line; Pyr.=Pyramidal tract C.N. = Caudate nucleus; Th. = Thalamus; Th.F. = Thalamo-frontal tract; F.P.=Fronto-pontine and Cortico-nuclear fibres; Putam. = Putamen; Gl.P.=Globus pallidus; L.N.=Lenticular nucleus; T.P.=Temporo-pontine fibres.

The direct pyramidal fibres are of some practical significance as they may be responsible in part for recovery of function after injury to the crossed pyramidal fibres; the direct pyramidal fibres do not as a rule extend beyond the lower cervical or midthoracic region. The pyramidal fibres finally end round short internuncial neurones (situated in the dorsal grey matter of the spinal cord) from which the impulses are relayed to the ventral horn cells.

Effects of Stimulating Motor Area. 1. Results in Apes.[1]—In the apes the motor area occupies the whole length of the precentral convolution; it extends backwards into the fissure of Rolando [sulcus centralis] (about one-third of the excitomotor area is "buried" in the fissure), but never

[1] Sherrington and Leyton, *Quart. J. exp. Physiol.*, 1917, *11*, 135.

on to the postcentral convolution. At its upper edge it extends down on to the medial surface of the hemisphere. Its anterior limit is variable and frequently indefinite; it may cross into the superior and inferior precentral fissures (Fig. 413). (In the ape, at least, it seems that the motor area does *not* correspond precisely to area 4.)

Electrical stimulation of this region gives rise to *coordinated* movements on the *opposite* side of the body. The law of *reciprocal innervation* holds good here: the antagonists relax simultaneously with the contraction of the protagonists. There are well-defined areas in the cortex which give rise to specific movements which are restricted to narrowly limited parts of the body; such restricted movements are often called "discrete" movements. Generally speaking, the body is represented upside down in the cortex, *i.e.* if stimulation is carried out in the excitomotor area from below upwards, bodily movements are produced in the following order: larynx, tongue, jaw, mouth, nose, eyelid, ear, neck, hand, wrist, elbow, shoulder, chest, abdomen, hip, knee, ankle, toes, perineal muscles. The hand is the part of the arm

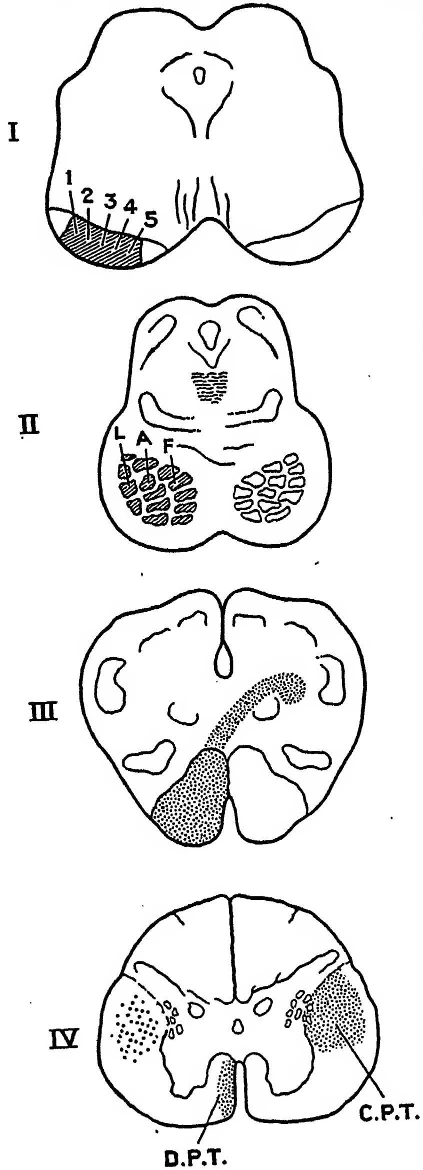


FIG. 412.—Course of Pyramidal Tracts in Ape. (After Leyton and Sherrington, *Quart. J. exp. Physiol.*, 1917, 77.)

I. Midbrain.
II. Pons.
III. Medulla (at level of decussation of pyramidal tracts).
IV. Spinal Cord (cervical region).
I. Stimulation of the points 5, 4, 3, 2, 1 in the midbrain gave rise to movements in the opposite side of face, neck, arm, trunk and leg respectively.
II. Stimulation of the points F, A, L, in pons produced movements in the opposite side of face, arm and leg respectively.
III, IV. To produce these sections, the motor cortex on the left side was exposed. The arm area was demarcated by stimulation and then excised. Time was given for the pyramidal "arm" fibres to degenerate. The animal was killed and sections of medulla and spinal cord were cut and stained by Marchi's method.
III. Note that the degenerate "arm" fibres are scattered throughout the pyramid in the medulla and are not topographically localized. Note also the beginning of the pyramidal decussation.
VI. Note that most of the pyramidal fibres have crossed to the opposite side (C.P.T. = crossed pyramidal tract); some remain ventrally on their original side (D.P.T. = direct pyramidal tract); a few fibres pass back to the lateral column of the original side.

which follows immediately on the face; this may perhaps be remembered by considering the close relation of the hand to the face in the taking of food. The hip follows immediately on the abdomen. The toes and the perineum are associated together on the highest part of the precentral gyrus and extend on the *medial* surface of the hemisphere.

The situation of these so-called " centres " (or " motor points ") for specific movements is not rigidly fixed anatomically. There is a variation in the distribution of the " centres " concerned with movement of specific parts both in different individuals of the same species and in the two hemispheres of the same individual. Furthermore, the response obtained from any cortical " motor point " is modified by the *previous excitation history of that point or of neighbouring or even of distant areas.* Repetition of a stimulus may give a larger response than on the first occasion (*facilitation*); or the response may be found to be *reversed, i.e.* the point which previously gave rise to flexion may now give rise to extension; or there may be *deviation* of the response, *i.e.* a movement is obtained from a different part of the body than on the former occasion. On the other hand, following stimulation of a motor point its excitability may *diminish,* reaching its minimum after seconds or minutes. These results indicate clearly that the motor areas possess a functional *flexibility* and that they are readily influenced by what is going on around them, probably because of the enormous wealth of neural associations between different regions. Prolonged faradization of any spot may give rise to widespread *epileptiform convulsions* (cf. p. 627). Cortical excitability is depressed by acidæmia and enhanced by a bout of overventilation (cf. p. 408).

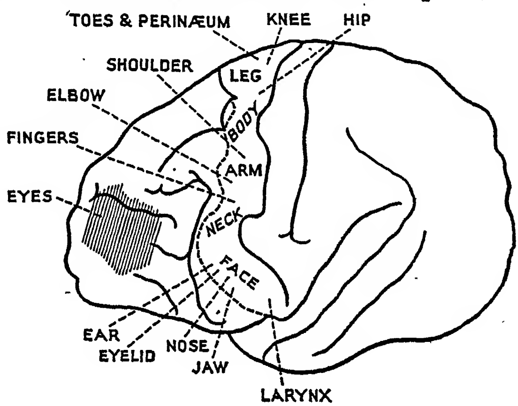


FIG. 413.—Diagram of Outer Surface of Cerebral Cortex of the Chimpanzee to show the Arrangement of the Motor Centres. (After Sherrington.)

Some further details may now be given: (i) *Tongue:* The deviation of the tongue may be to the same or the opposite side, and all kinds of complex contortions may be produced. (ii) *Jaws:* When the symphysis is split the response is mainly contralateral. (iii) *Vocal Cords:* The response is bilateral to some extent. (iv) The only *eye movement* represented in the ascending frontal convolution is eyelid closure. (v) *Hand:* The thumb is represented lowest in the hand area and the fingers highest. Isolated responses may be obtained from the thumb or the index finger; occasionally the thumb and index respond together as, for instance, by abduction of the thumb and extension of the index to produce a kind of " let-go " movement. In the case of other digits isolated movements are rare. (vi) *Leg:* The leg centres are more commingled than the arm centres. (vii) *Anus:* Anal protrusion is usually symmetrical, but sometimes even this movement may be mainly contralateral.

Stimulation of a motor point is not identical with stimulating the cells of origin of the pyramidal tract; the whole depth of the grey matter in the vicinity of the motor point may be stimulated, including the four layers external to layer V, layers which can be regarded as *massed internuncial neurones* which receive impulses from adjacent and more distant areas of the cerebral cortex and from lower levels, *e.g.* the cerebellum. The wealth of the connections within any cortical area and between cortical areas is emphasized in Fig. 396 and on p. 616. The varying activity of this internuncial background will largely determine, in the case of any stimulation experiment, which cells of layer V discharge, which pyramidal tract fibres convey impulses and consequently which lower motor neurones are activated. It is readily understandable that the results of cortical stimulation may vary with the "previous history" of the cerebral cortex; in fact, it is surprising that approximately similar movement-patterns result from cortical stimulation at regular intervals. It must also be remembered that the corticospinal tracts do not end directly round ventral horn cells, but connect mainly with those dorsal horn internuncials which also receive many connections from incoming dorsal nerve roots. The effects produced by corticospinal impulses in this spinal cord region will depend largely on the "background activity" induced in it by afferent impulses from muscles and skin. In view of the relative crudity of the experimental conditions it becomes a matter for wonder that the stimulation of motor points produces movements which are so well co-ordinated, display perfect reciprocal innervation and, as mentioned below, are so suggestive of "fragments" of voluntary movement.

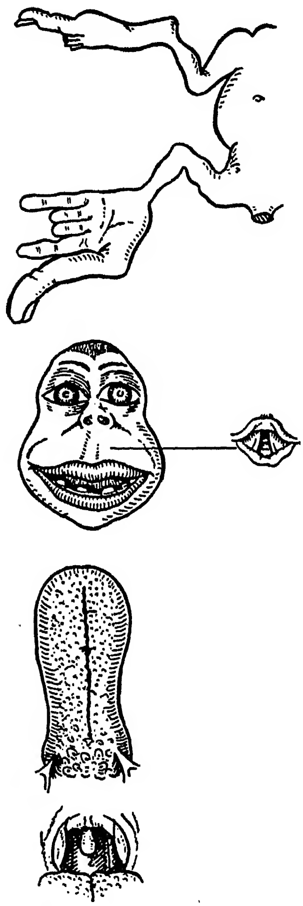


Fig. 414—Diagrammatic Homunculus to illustrate the Relative Position and Size of the Motor Areas for the different parts of the body in Man. (After Penfield.) (To economize space only one arm, half the trunk, and one leg are shown.)

An enormous variety of responses can be obtained from stimulation of the excitomotor area in the ape. Sherrington, for example, has recorded more than four hundred types of movement-pattern produced in this way. It is important to emphasize that though the movements thus obtained are perfectly co-ordinated they are not in themselves *complete* acts which could effectively fulfil some set purpose; rather are they fragmentary or fractional in nature, or *unitary components of a larger and more complex whole*. Their functionally incomplete character is most readily detected in the unilateral nature of the response even in parts where the two sides invariably work

together, *e.g.* the jaws, the vocal cords or (most convincingly) the anus. The fact that under natural conditions the movements of these parts are invariably symmetrical and bilateral proves that the *motor cortices on the two sides of the brain are very closely linked together*. As the movement-fraction represented by each motor point is so small, the motor cortex is endowed with great flexibility because these motor points can be combined together in an immense variety of permutations and combinations to carry out almost an infinite variety of purposive acts. A useful analogy can be drawn between this feature of the motor points and the amino-acids which are supplied to the tissues; a relatively small number of chemical units can similarly be built up into an infinite variety of specific proteins.

2. **Results in Man.**[1]—In patients under local anæsthesia, stimulation of the motor area elicits responses closely resembling those described for the ape, *i.e. discrete isolated* movements (on the opposite side) of a single segment of an extremity or a single part of the trunk or head. The cortical representation is arranged as in the ape, except that separate foci exist for each of the fingers, and these occupy a relatively large area. The focus for the thumb is most inferior, and that for the fifth finger most superior. Stimulation of the points controlling the upper part of the face, the pharynx, the vocal cords, and the muscles for closing the jaws, usually gives *bilateral* reactions. Stimulation of the neck area usually causes the head to turn to the opposite side. Electrical stimulation of the most inferior part of the precentral gyrus produces rhythmic co-ordinated *movements of the lips, tongue, mandible, larynx, and pharynx*. Epileptic attacks beginning in this area commence with the same type of movement—chewing, licking, swallowing, and grunting. Fig. 414 shows in very striking fashion the relative position and size of the motor areas for different parts of the body in man.

Jacksonian Fit.—A fit which arises in a localized part of the cortex is known after Hughlings Jackson as a "Jacksonian fit." If the fit spreads it will involve successively regions of the body supplied by related parts of the affected cortical area. Thus a fit beginning in the thumb will spread via the wrist, elbow, and shoulder, to the trunk, and via the neck to the face (cf. Fig. 413).

Effects of Stimulating Cortical Eyefields (Fig. 407).—These are found mainly in the middle frontal convolution in area 8 (*frontal eyefield*)[2] and in the occipital lobe in areas 17 and 18 (*occipital eyefield*). The response usually involves both eyes and the movement most commonly obtained is a conjugate deviation of the eyes to the opposite side, *e.g.* stimulation of the left cortex causes both eyeballs to move to the right (Fig. 415). This movement may be accompanied by downward and upward deviation; very rarely convergence of the eyes may be observed. Eye-opening is also associated with this region. *Autonomic* responses affecting the eyes may also be elicited, *e.g.* dilatation of the pupils, secretion of tears.

Stimulation of the *sixth nucleus* also causes conjugate deviation of both eyes—this time to the side of the stimulated nucleus; *e.g.* stimulation of the right sixth nucleus causes conjugate deviation of both eyes to the right owing to the simultaneous contraction of the right lateral rectus and the left medial

[1] Penfield and Rasmussen, *Cerebral Cortex of Man*, N.Y., 1950.

[2] In man, eyeball movements often result also from stimulating points in the precentral gyrus itself.

rectus. The exact innervation of the medial rectus is uncertain. It is suggested that internuncial fibres establish connection between the (right) sixth nucleus and the (right) third nucleus (Fig. 415), whence fibres pass (perhaps via the opposite third nucleus) to the opposite (left) third nerve and thence to the opposite (left) medial rectus.

The conjugate deviation commonly resulting from stimulating the frontal eyefield can now be explained. The (left) cortico-nuclear fibres cross the midline in the midbrain to supply the opposite (right) sixth nucleus which is

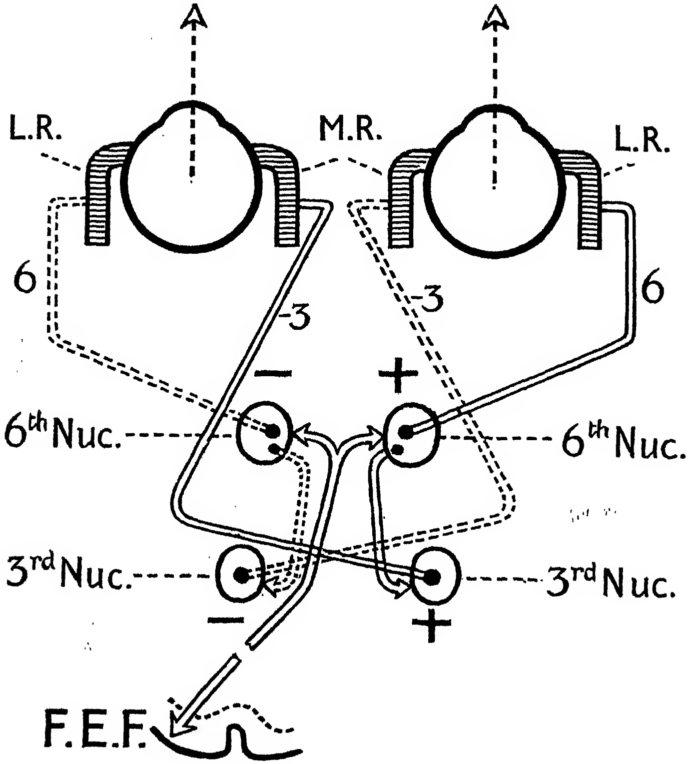


Fig. 415.—Cortical Control of Lateral Movements of the Eyeballs.
F.E.F.: frontal eyefield; L.R., M.R.: lateral, medial rectus muscles; 3, 6: 3rd and 6th nerves. +, –: excitation or inhibition of nucleus. The diagram has been simplified and shows the third nucleus directly supplying the contralateral medial rectus. (Diagram by Professor David Slome.)

stimulated with resulting contraction of the (right) lateral rectus; via the internuncials between the (right) sixth and third nucleus, the *left* third nucleus, and the *left* third nerve, the left medial rectus is stimulated.

Irritative lesions of the frontal eyefield or of the sixth nucleus in man produce results like those of experimental stimulation in apes. A destructive lesion of the (say, right) sixth nucleus leaves the opposite (left) sixth nucleus unopposed and produces, as expected, conjugate deviation of both eyes away from the side of the lesion (*i.e.* to the left).

Reciprocal Innervation.—Sherrington has demonstrated that the law

of reciprocal innervation which applies to reflex action (p. 545), voluntary movements (p. 649), and movements elicited from stimulating the motor area (p. 634), also applies to the cortical control of lateral movements of the eyeballs; *i.e.* the impulses from the frontal eyefield that stimulate the contralateral sixth nucleus reciprocally inhibit the ipsilateral sixth nucleus, this latter inhibition being an integral part of the total response (Fig. 415). The experiment is carried out as follows:

In the monkey, if all the nerves supplying the muscles of the eye are cut, the tension of the connections of the eyeball keeps the globes parallel. Cut the third and fourth nerves of the left eye. The only nerve left intact is the sixth, and naturally the eye rotates outwards from the unopposed action of the lateral rectus (Fig. 416, B). Expose the frontal lobes on the left side, *i.e.* on the same side as the eye operated on. Stimulation of the left frontal eyefield produces contraction of the right lateral rectus and of the left medial rectus. In the left eye the medial rectus cannot contract, because the nerve carrying its motor fibres (third) has been cut. It is obvious that the medial and lateral recti are antagonistic muscles, and that if one were stimulated the other should be inhibited. In the case of the left eye, as the motor cells supplying the medial rectus are stimulated from the cortex, the cells innervating the lateral rectus should be inhibited. This is exactly what happens. The lateral rectus relaxes, and the left eye moves to the middle line (Fig. 416, C). Note that in this experiment the antagonist centre has been inhibited (*i.e.* the cells supplying the lateral rectus) although no contraction of the protagonist (medial rectus) has actually taken place (because the nerve to it was cut). The pattern of reciprocal innervation in this case thus depends on appropriate impulses coming down to the cranial nuclei from the cerebral cortex.

Results of Destruction of Motor Area or Pyramidal Tracts.—(1) Extirpation of Motor Areas.—Sherrington carried out careful extirpation experiments in apes. The left motor cortex was exposed and the arm area demarcated by electrical stimulation. The region yielding primary movements of fingers, wrist, and elbow was excised. A few hours later the opposite (right) arm showed drooping of the wrist, weakness of the elbow, and to a less extent of the shoulder; no movement was possible in the fingers. One month later good *recovery* of most arm movements had occurred, but there was lasting *decrease in the skill with which the fingers could be used.* The operation area was re-exposed, but on electrical stimulation it yielded no response. A little more of the adjacent cortex was excised; some transient weakness of the shoulder developed, but the distal part of the limb was unaffected. Some months later the opposite (right) arm area was excised; this had no appreciable effect on the originally paralysed (right) arm.

(2) Section of Pyramidal Tracts.—Removal of the arm area destroys the cells of origin not only of the corticospinal fibres but of many extrapyramidal fibres also. To determine the rôle of the pyramidal tracts alone, the tracts must be cut at a level where they are not commingled with other descending tracts which subserve other functions. The medullary pyramid may possibly consist almost exclusively of corticospinal fibres; should this view prove correct then *section of the medullary pyramid* produces the uncomplicated results of loss of pyramidal tract function.

(i) *Effects on Movement.*—In rhesus monkeys unilateral section of the

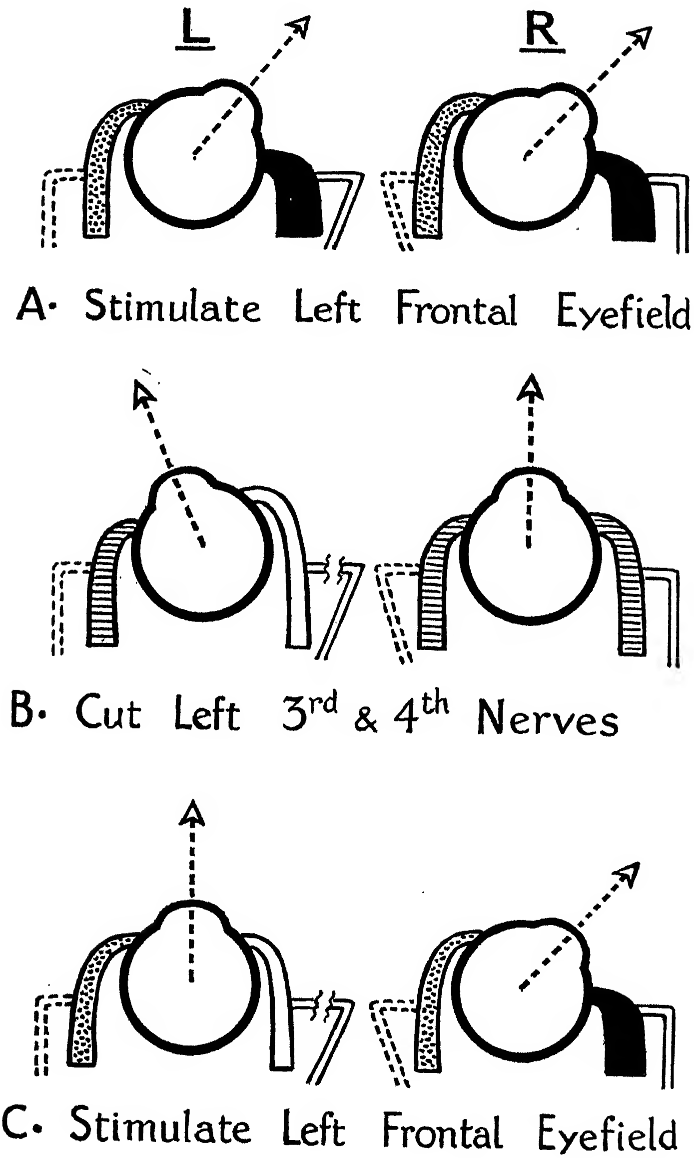


FIG. 416.—Sherrington's Demonstration of Reciprocal Innervation of Eye Muscles.

A. Stimulate left frontal eyefield conjugate deviation of both eyes to the right.
B. Cut left third and fourth nerves; deviation of left eye to left.
C. Repeat stimulation of left frontal eyefield; right eye deviates normally to the right and left eye returns to midline. (Diagram by Professor David Slome.)

medullary pyramid produces disturbances of movement closely resembling those resulting from extirpation of the motor area. Thus there is marked weakness (not complete paralysis) of the opposite arm and leg; the hand is most gravely affected, discrete movements of the fingers being lost. Simpler movements persist, *e.g.* striking, kicking, scratching, reaching, grasping, walking, turning the body towards external objects; but these movements, also, are less easily and accurately performed; they seem to involve great effort, are weak and tremulous, fatigue readily, and are accompanied by signs of great "irritation" on the part of the animal. The symptoms persist unchanged for years.

(ii) *Effects on Tone and Reflexes.*—There is a *decrease* in muscle tone in the limbs (tone in the back, neck, and thorax is decreased very little). The deep reflexes are slow but not decreased in size. The superficial reflexes (*e.g.* abdominal, cremasteric) are abolished at first; later they return but are difficult to elicit, sluggish and weak.

Similar changes in tone, reflexes, and movement result from section of the crossed pyramidal tract in the lateral column of the spinal cord.

A sharp contrast thus exists between the *decreased muscle tone and reflexes following allegedly pure section of the pyramidal tract* and the *spasticity and exaggerated reflexes following extirpation of the suppressor band 4s* (p. 624), or clinical lesions of the pyramidal tract (p. 642).

In the chimpanzee, two "release" phenomena develop after section of the medullary pyramid: (*a*) the *Babinski plantar response* makes its appearance (cf. p. 645); (*b*) the proprioceptive *grasp reflex* can now be obtained: if an object is firmly pressed into the palm, it is reflexly grasped.

Rôle of Pyramidal Tracts in Voluntary Movement.—The two outstanding facts described: (i) that artificial stimulation of motor points produces movements which resemble fragments of a voluntary movement, and (ii) that impairment or loss of voluntary movements results from extirpation of the motor cortex or section of the pyramidal tracts, indicate that the pyramidal tract is a path which is employed in voluntary movement and in fact is indispensable for the performance of certain types of movement. This statement does not mean that the motor cortex initiates voluntary movements. The pyramidal tracts may be regarded as a long *internuncial* pathway, linking the great afferent pathways which end in the cerebral cortex (especially those from the "distance receptors" (eyes, ears), and from the muscles (directly and via the cerebellum) with the lower motor neurones (cranial and spinal). The regulation of the discharge of the motor cortex to produce effective purposive movements, "correctly ordered in space and time," is a matter which is ill understood; it is considered, however, on pp. 646 *et seq.*

Clinical Lesions of the Pyramidal Tracts.—In the past the significance of the multiplicity of the descending pathways from the cerebral cortex was not fully recognized nor was the fact that throughout most of its course the pyramidal tract is commingled with other descending tracts. In man, injury to the pyramidal tract in the brain gives rise to weakness or paralysis on the opposite side of the body; the condition is called *hemiplegia.* The symptoms and signs in any individual case depend on the severity of the lesion, whether it is acute or chronic, and on the extent to which the other descending (extrapyramidal) pathways are involved. In view of the

experimental evidence presented above, the disturbances of voluntary movement which occur in hemiplegia are attributed to loss of the pyramidal influence. Commonly the changes in tone and reflexes in these patients resemble those resulting from extirpation of area 4s in animals; these changes in man are explained on similar lines, *i.e.* they are attributed to injury to the descending inhibitory (extrapyramidal) pathway. As might be expected, there is no constant relationship between the severity of loss of willed movements and the degree of hypertonus present.

Hemiplegia.—1. Acute Lesions in Man.—An *acute* lesion in the internal capsule, where the pyramidal tracts are condensed in a small area, produces initially a *stage of shock* in which all the muscles of the opposite side are toneless and in which no reflex movements can be elicited; in these respects, the condition resembles spinal shock (p. 690). These findings again emphasize the degree to which the higher levels of the nervous system participate in so-called spinal and brain stem reflexes. Some weeks later reflex activities return to the affected side of the body, showing that such reflexes can be mediated by the isolated lower levels. A period of "adjustment," however, has been necessary, indicating that the lower and higher nervous levels are normally closely integrated to function as a co-ordinated whole.

The paralysis on the affected side affects voluntary movements of the face, leg, and arm: Movements which usually involve both sides of the body, *e.g.* those of respiration, movements of the back, and abdominal wall, are generally retained; it would seem that one pyramidal tract can control the lower motor nuclei supplying the muscles concerned, on both sides. If the cortico-nuclear fibres escape injury, eyeball movements persist. *Emotional movements remain intact*; frequently they are elicited very *readily* and are *exaggerated* in character. This last finding demonstrates two points: (i) emotional movements are mediated by non-pyramidal pathways and are *inhibited by the pyramidal tract*; (ii) the paralysis in pyramidal lesions is one of *movements* and not of *muscles.* Thus the patient cannot voluntarily wrinkle up his forehead, frown or whistle, but the muscles concerned in these acts are employed perfectly efficiently when the patient's face exteriorizes pleasure, surprise, or annoyance (cf. p. 663).

2. Stage of Recovery. Chronic Lesions.—As the stage of shock passes away, or in chronic lesions, the clinical picture as explained is that of paralysis and disturbances of tone and reflexes. The common clinical syndrome which is described below is attributable to a *combined lesion of the pyramidal tract and the descending inhibitory path from the suppressor band 4s.*

(1) Muscle Tone and Posture.—Muscle tone is abnormal in distribution and excessive in degree in certain muscles, *i.e.* the limbs are placed in an abnormal position and tend to be fixed there; to this latter feature the term *spasticity* is applied. The *upper limb,* instead of lying at the side in the customary way, is adducted at the shoulder, the elbow is semi-flexed, the forearm is pronated, and the wrist and fingers are flexed. The limb is involuntarily maintained in this unnatural position indefinitely without fatigue; it can only be moved passively with difficulty. The *leg* is adducted, extended at the knee and physiologically extended at the ankle (anatomical plantar flexion).[1] The muscles do not waste greatly, because though not used in voluntary movement they are continuously in action to maintain the posture

[1] Cf. Fig. 372, p. 591.

described; the electrical responses are normal. Lesions of the dorsal nerve roots abolish the spasticity and the muscles become flaccid.

The limbs show the *shortening* and *lengthening* reaction. The latter is demonstrated as follows: If one attempts to flex the knee (with the patient supine) the initial 5°–10° of flexion can be carried out without much difficulty; beyond this point marked resistance is encountered, because the stretching force applied to the quadriceps elicits a reflex contraction (cf. p. 558); if great force is applied, the quadriceps suddenly relaxes, and the knee rapidly passes into the flexed position—this is the lengthening reaction set up by excessive stretch which stimulates the self-inhibiting afferents in the muscle (cf. p. 586). The lengthening reaction is described by the clinician as *clasp-knife rigidity*.

(2) Associated Reactions.[1]—This term is used to describe certain movements which can be reflexly aroused on the affected side by such "semi-involuntary" movements as yawning and stretching and by any forceful

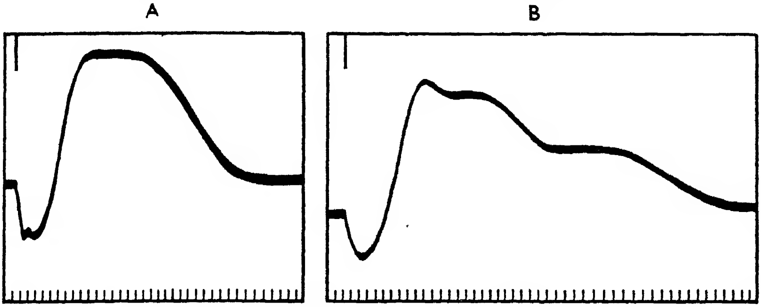


Fig. 417.—Knee-jerk in Normal Man and after Lesion of Pyramidal Tract.

The signal (vertical line above) indicates the application of the stretch stimulus which produces a sudden depression of the leg as indicated by the initial descent of the record. Time in 0.01 sec.

A. Normal response. Note sharp ascent, plateau, and more gradual but smooth descent.
B. Knee-jerk following pyramidal lesion. The characteristic feature is the much slower relaxation which shows a "hump" on its course. (Blake-Pritchard and Walshe, *Lancet*, 1929.)

sustained voluntary muscular contraction on the normal side; the responses are modified by the position of the head relative to the trunk, *i.e.* by *neck reflexes*. Thus in the standing hemiplegic subject, looking *straight ahead*, when the normal fist is clenched, the spastic arm moves slowly into increased flexion at elbow, wrist, and digits. If the procedure is repeated with the head rotated *towards the hemiplegic side*, the spastic arm moves into extension and abduction, the fist remaining closed. Repeat the procedure again with the head rotated *towards the sound side*: the spastic arm moves into increased adduction and flexion. In other words, the jaw limbs extend and the vertex limbs flex. In some cases, active rotation of the head against resistance is sufficient to induce similar postural changes in the spastic limb.

(3) Deep Reflexes.—The deep (*i.e.* tendon) reflexes are modified in a characteristic way. The ones commonly elicited are the *knee-jerk* and *ankle-jerk* in the lower limb, and the *triceps* and *supinator* jerk in the upper limb. The deep reflexes are simply fractionated stretch reflexes, and their general characteristics can be studied in the case of the human knee-jerk.

[1] Usually referred to clinically as associated movements.

Knee-jerk.—With the knee supported in the flexed position, the patellar tendon is sharply tapped ; the muscle is stretched, appropriate sensory nerve endings are stimulated (muscle spindles, tendon organs) and impulses pass up into the third and fourth lumbar dorsal nerve roots to enter the spinal cord and end directly round ventral horn cells. Efferent impulses pass out in the lumbar ventral nerve roots to produce contraction of the quadriceps extensor and at the same time reciprocal inhibition of the antagonistic hamstrings takes place. The knee-jerk is abolished when any part of this reflex arc is destroyed. It should be borne in mind then when we studied the stretch reflexes experimentally we were dealing with *sustained* stretch of muscles which were contracting *isometrically* ; in the case of the human deep reflexes we are observing the response to a *transient* stretch of a muscle contracting *isotonically*. A graphic record of the normal knee-jerk shows that there is a brisk contraction followed by a plateau and a gradual relaxation (Fig. 417, A) ; the leg falls slowly and deliberately back to its original position. The ratio of the rise to the fall may be 1 to 2. The delayed relaxation (or persistent contraction) is due partly to after-discharge (cf. p. 539) of the ventral horn cells and partly to an appended shortening reaction (cf. p. 586

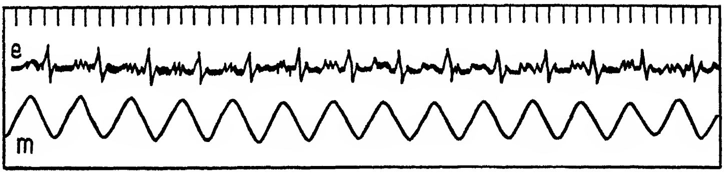


Fig. 418.—Clonic Response of Quadriceps to Stretch.

Electrical (*e*) and Mechanical (*m*) Records. Clonic response of quadriceps muscle (at 13 a second) elicited by attaching the muscle to an isotonic lever weighing one kg., and so subjecting it to stretch. Note the rhythmic character of the electrical record. (Sherrington and others, *Reflex Activity of the Spinal Cord*, 1932.)

and Fig. 369), *i.e.* as the knee begins to flex again the quadriceps is subjected to renewed stretch and responds by partial contraction.

The deep reflexes like the stretch reflexes vary with the degree of muscle tone initially present and with the amount of nervous system which is intact. In hemiplegia there is commonly increased extensor tone. The knee-jerk (like the other deep reflexes) is *exaggerated*, and what is even more characteristic, it is more *sustained*, *i.e.* relaxation is still further prolonged (Fig. 417, B). Not infrequently there is a secondary contraction or "hump" during relaxation.

Ankle clonus is often present, and is also due to the state of heightened muscle tone. When the ankle is forcibly dorsiflexed, rhythmic movements of dorsiflexion and plantar flexion take place at the ankle-joint. Owing to the increased tone of the posterior calf muscles, stretching of the fibres sets up not a simple reflex contraction or ankle-jerk, but a rhythmic series of contractions or clonus. In animal experiments also it is sometimes found that stretch of a muscle contracting *isotonically* will elicit a rhythmic response instead of the usual sustained contraction (Fig. 418). This is doubtless due to a *synchronous* (instead of the more usual asynchronous) discharge of the ventral horn cells, leading to a tremulous partial tetanus (cf. p. 501).

Sometimes a clonic response is obtained on tapping the patellar tendon (*patellar clonus*).

The changes in muscle tone, posture, and deep reflexes described above must, in view of the experimental evidence, be attributed not to a lesion of the pyramidal (*i.e.* corticospinal) tract, but to interruption of the inhibitory fibres from the suppressor areas (mainly 4s) to the nuclei of the reticular formation in the brain stem and thence to the spinal cord. The changes described represent a series of " positive " release phenomena ; *i.e.* removal of the cortical inhibitory influence releases lower levels to display their inherent pattern of activity. The resemblances between the changes in tone and reflexes in human hemiplegia and in decerebrate rigidity in the cat are obvious ; the only notable difference is in the posture of the fore limbs (arm in the case of the human hemiplegic). The amount of cerebral damage needed to produce the so-called " spasticity " in man and the equivalent, so-called " rigidity " in the cat differs in the two species ; in the cat all the brain must be cut off above the mid-collicular level (in the midbrain) ; in man a lesion of the cortical inhibitory pathway sparing the rest of the nervous system is probably adequate.

(4) Superficial Reflexes.—The superficial reflexes commonly tested for are :

(i) *Abdominal* : stroking the skin of the abdomen produces contraction of the underlying muscle.

(ii) *Cremasteric* : stroking the inner side of the thigh results in the testis being retracted towards the inguinal canal.

(iii) *Plantar* : stroking the sole of the foot produces a downward movement (plantar-flexion) of the great toe and the small toes.

All the superficial reflexes are *lost* on the affected side in cases of hemiplegia. As these reflexes are also lost when the medullary pyramid is cut, they are mediated by the pyramidal tract ; thus their loss in hemiplegia is due to the injury to the pyramidal tract.

(iv) *Babinski response* : in hemiplegia an abnormal reflex can be elicited from the sole of the foot, different in every respect from the usual plantar response. The stimulus employed must be *nocuous* or painful, *e.g.* firm scratching with the finger-nail. The reflex can be elicited from a fairly wide area, from the calf, the thigh, and even as high as the groin ; it consists first of an *upward* movement (dorsiflexion) of the great toe and fanning out of the small toes. The anatomists misleadingly call dorsiflexion of the great toe " extension." The abnormal toe response is therefore sometimes called the " extensor response " ; this term should not be used and the response should be described (after its discoverer) as the Babinski sign or the Babinski response. The Babinski response is a " fraction " of a released spinal flexor or withdrawal reflex. Thus, in some cases of hemiplegia, nocuous stimulation of the sole of the foot produces a reaction extending far beyond the toes ; there may be dorsiflexion (upward movement) of the ankle, flexion of the knee and even flexion of the hip. The more extensive reaction is identical in every way with the spinal flexor reflex (p. 692). The Babinski response is obviously a release phenomenon ; as it appears experimentally after section of the medullary pyramid its occurrence in man must be attributed to a lesion of the pyramidal tract.

(5) Disturbances in Voluntary Movement.—In a recovering case of

hemiplegia, considerable improvement occurs in the leg and the patient may walk about with little more disturbance than a slight limp; a good deal of power returns to the arm and face. The incidence of *permanent* loss of movement, in order of severity, is as follows: movements of the hands and especially the digits; extension of wrist; supination of the forearm; abduction and elevation of the upper arm; dorsiflexion of the foot and toes; flexion of the proximal leg joints. The recovery may be attributed to restoration of function to temporarily damaged pyramidal fibres or to the more effective use of extrapyramidal pathways arising in the cortex.[1]

Walshe has analysed carefully the loss of movement in the arm and hand in *progressively developing hemiplegia.* Paresis first appears in the movements of the hand and digits and spreads so as to involve the proximal part of the limb. The first movements to disappear are those in which the interossei, lumbricales, and the flexors and opposers of the thumb are involved. Abduction and extension of the fingers are affected before adduction. Certain combinations of movements disappear, as is well shown by asking the patient to touch each finger-tip in succession with the tip of the thumb; there is a stage in the development of the hemiplegia in which, while each digit can be flexed, the necessary movements of adduction and apposition are lost and the thumb and finger-tips are not effectively approximated "but flex futilely into the palm." The limb is progressively "denuded of movements" and therefore of movement combinations; as the number of retained movements dwindles, the surviving patterns of motion are simple and few, and qualify for the term "stereotyped movements."

Bilateral Pyramidal Lesions.—Bilateral section of the medullary pyramids in monkeys produces complete motor helplessness. Extensive bilateral lesions in man may likewise produce loss of all movements in the affected parts. Thus, in a patient in whom the pyramidal fibres controlling movements of head, neck, and thorax were bilaterally destroyed, all voluntary movements of the facial muscles were lost; biting, swallowing, elevation of the palate and laryngeal movements could not be carried out. Although automatic rhythmic breathing continued, forced inspiration or expiration or voluntary coughing could no longer be carried out. *Emotional exteriorization was vigorous* and was produced at the slightest provocation (cf. p. 663).

The rôle of the pyramidal system in willed movements is considered further below.

GENERAL SURVEY OF VOLUNTARY MOVEMENT. SPEECH

Voluntary Movements.[2]—Physiological knowledge is still insufficient to enable us to describe in detail what happens in the central nervous system during the execution of a voluntary movement. To give only one example, the efferent pathways arising from the precentral cortex are so complex (Fig. 408) that it is hard to conceive how their activities are harmoniously co-ordinated to produce efficient muscular action. There is no doubt that voluntary movements are impaired whenever there is an interruption of the afferent channels to the brain, especially those from the sense organs (pro-

[1] Cf., however, Lassek, *Brain*, 1950, *73*, 95.

[2] Beevor, *Croonian Lectures (for 1903) on Muscular Movements*, London, 1904; new edn., 1951. Wilson, *Lancet*, 1925, ii, 1 *et seq.*

prioceptive from the muscles and labyrinth and from the cerebellum, exteroceptive from the skin and eyes). What happens between the arrival of the impulses in the receptor areas of the cortex and the discharge of efferent impulses from the precentral motor cortex is unknown, and therefore much written about.[1] The process of learning, by which movements which are initially executed clumsily and with difficulty become ultimately skilful and easy, has not yet been accounted for in physiological terms; nor is anything known about the relationship between the mind or will and the brain. Some physiologists use the word brain as though it were synonymous with mind; such usage is confusing. Other physiologists try to avoid using the concept of mind altogether; such timidity does not promote clarity of thought or exposition.

Voluntary movements are voluntary in their *aim*, not in their *means*. For

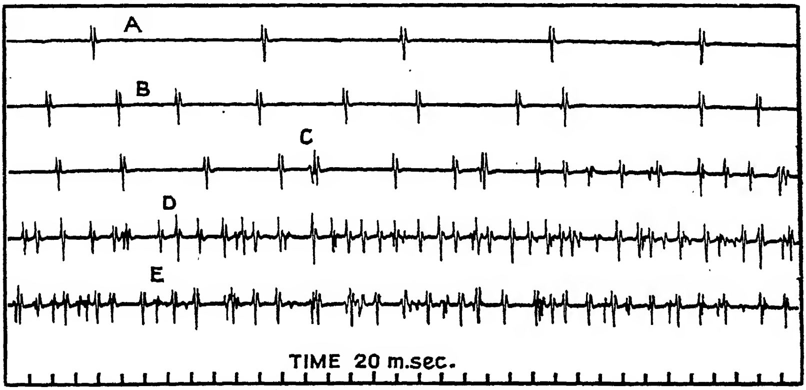


FIG. 419.—Muscle Action Potentials during Voluntary Movement.

Action potentials (recorded with concentric needle electrodes) in the first dorsal interosseous muscle The motor unit potentials are triphasic as is not uncommonly the case. Records are continuous from left to right and from above downwards. Time, 20 m.sec.
A. Commencing contraction potentials are at low frequency.
B. Frequency increasing.
C. Additional motor units come into action.
D, E. Discharge frequency of active units rises.

simple experimental and clinical tests, the aim is usually a displacement of joints or of segments of the body in a certain direction and to a certain extent; the subject is told to clench the fist, flex the elbow, extend the knee and the like. He and we judge by inspection, how well he has succeeded in carrying out the instructions. The subject is not consciously concerned with the events in the nervous system and muscles which cause the movement; he wills "bend the arm," and it happens; visual and proprioceptive impulses tell him that it has happened. In everyday activities, our aim is generally not directed to parts of the body but to external objects; we cut the bread, open the door or kick a ball; again we can easily judge whether we have succeeded or failed.

An attempt is often made to draw a distinction between "voluntary" and "automatic" movements, the movement of walking frequently being

[1] It should be remembered that some common movements are not represented in the motor cortex, *e.g.* walking, scratching, swimming.

placed in the latter category. But no really hard-and-fast distinctions can be made. Walking is learnt by the child with difficulty, and at first requires constant attention. With the passage of years and increased practice it becomes less " voluntary " and more " automatic " in character, so that we can walk about while thinking of something quite different. It is preferable to grade movements into " less or more automatic " or " less or more voluntary," instead of attempting to divide them sharply into voluntary and automatic movements.

Many groups of muscles play a part in carrying out even quite a simple voluntary movement. We may recognise :

(1) Prime Movers or Protagonists.—For example, the chief protagonists in flexion of the elbow are the biceps, brachialis and brachioradialis. The ventral horn cells supplying each of these flexors are called its *motor neurone pool*. The extent and strength of a voluntary movement (as of a reflex movement) are graded in two ways :

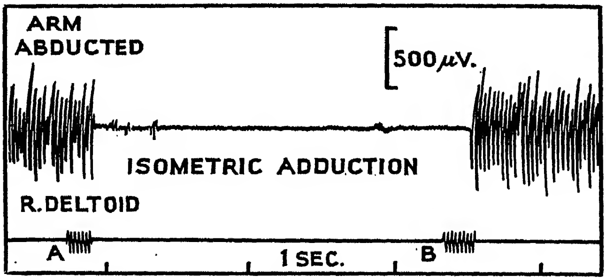


Fig. 420.—Reciprocal Inhibition of Deltoid in Man. (From an experiment by W. F. Floyd and P. H. S. Silver.)

Muscle action potentials recorded with surface electrodes.
At the beginning of the record the arm was abducted against gravity ; there is marked electrical activity in the deltoid.
At first signal, A, attempt to abduct arm against a resistance which prevents movement from taking place (isometric adduction). The deltoid is reciprocally inhibited ; the action potentials disappear.
At second signal, B, attempt to abduct arm against resistance ; electrical activity in the deltoid reappears.

(i) By varying the *number* of motor neurones and therefore, the number of motor units in action.

(ii) By varying the *discharge rate* of the individual motor neurone ; as the frequency of the discharge increases, the weak tremulous sub-tetanus is converted into the strong sustained full tetanus.

These facts can be well demonstrated by using the concentric needle electrode technique. When a muscle is completely relaxed in man, action potentials are absent. As the movement begins (Fig. 419), action potentials appear ; the frequency is low at first and gradually increases. As the contraction becomes stronger, additional motor units come into action (Fig. 419, C). When the contraction becomes maximal, the electrical record shows that many units are in action and are discharging at high frequency (60–100 per sec.) ; the potentials are irregularly spaced and vary in shape and size (Fig. 419, D, E).

Elbow-flexion, under different circumstances may involve more activity in one of the flexors and less in another, *i.e.* different parts of the flexor motor

neurone pool may be used; another way of expressing this idea, is to say that the flexor motor neurone pool can be *fractionated* in different ways.

(2) ANTAGONISTS.—These are the muscles whose contraction or elastic resistance would interfere with the desired joint displacement. The antagonists always act in a manner which facilitates the execution of the desired movement (or the maintenance of the desired posture). In different sets of circumstances the antagonists may react in one of the following ways:

(*a*) Undergo reciprocal inhibition.

(*b*) Contract together with the protagonists (co-contraction).

(*c*) When the movement takes place in the direction of gravity they (the antagonists) contract, but the protagonists which would act in the same direction as gravity, do not.

(i) *Reciprocal Inhibition of Antagonists.*—The antagonists respond in this way during *movements against resistance.* Reciprocal inhibition occurs during voluntary movements of the eyeballs in the monkey (p. 638) and in movements experimentally elicited from the motor cortex (p. 634). It can be readily demonstrated during voluntary movements in man.

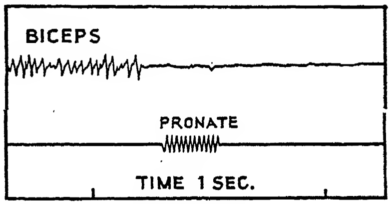


FIG. 421.—Reciprocal Inhibition of Biceps during Pronation in Man. (From an experiment by W. F. Floyd and P. H. S. Silver.)

Action potentials of biceps recorded with surface electrodes.
At the beginning of the record the elbow is flexed against gravity with the forearm in supination. There is marked electrical activity in the biceps.
At the signal, the forearm is *pronated* forcibly, without moving the elbow. The biceps is reciprocally inhibited. It is felt to relax and its electrical activity ceases.

(*a*) The arm (extended at the elbow) is abducted in the coronal plane to a right angle, *i.e.* it is raised to the level of the shoulder. The deltoid (the protagonist) can be felt to be strongly contracted; using surface electrodes numerous muscle action potentials are recorded (Fig. 420). The subject is then told to try to adduct the arm (*i.e.* lower it) against the resistance of the observer's hand which he cannot overcome. No movement of the subject's arm takes place and it remains in the abducted position. But during the attempt the deltoid has become the antagonist and is, therefore, reciprocally inhibited; it becomes quite soft and the action potentials temporarily disappear.

(*b*) The elbow is flexed against gravity with the forearm in the supinated position; all the elbow flexors contract. The biceps, however, contracts for two reasons: it not only flexes the elbow, but also supinates the forearm. The subject is now told to *pronate* the forearm forcibly without moving the elbow. At this stage of the experiment, the biceps motor neurone pool is receiving antagonistic impulses: excitatory, to maintain elbow-flexion, and inhibitory, to facilitate pronation of the forearm. On balance the biceps is reciprocally *inhibited* because its action opposes pronation (Fig. 421). The brachialis, which is purely an elbow-flexor, remains contracted.

(ii) *Co-contraction of the Antagonists.*—Sometimes the antagonists may cooperate with the protagonists by contracting too, as in the straight left of the boxer, in which protagonists and antagonists are both simultaneously contracted to keep the joint unmoved.[1]

[1] Cf. also Positive Supporting Reaction, p. 589.

(iii) *Contraction of Antagonists.*—When the force of gravity is the protagonist the antagonists may *reciprocally contract* to oppose gravity.

(*a*) As already explained, during comfortable standing in man the erector spinae and the anterior abdominal muscles show little activity (Fig. 422). On leaning slightly backwards the erector spinae contracts momentarily ; gravity then comes into action and would cause the subject to fall backwards. It is resisted by the action of the rectus abdominis which contracts forcibly and thus either arrests the backward movement, or enables it to proceed smoothly and in a properly graduated manner. Conversely on leaning forwards, the

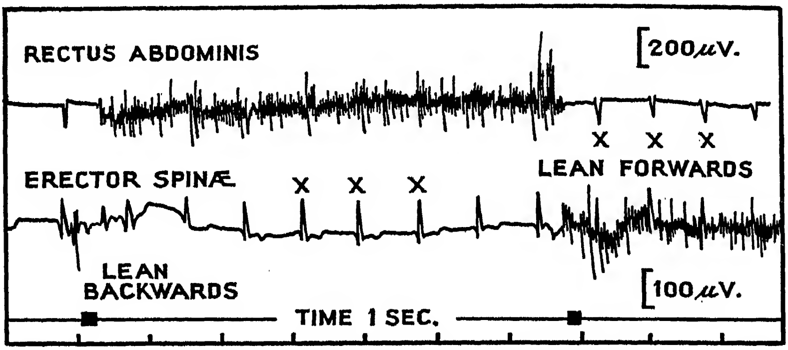


FIG. 422.—Pattern of Activity of Erector Spinae and Rectus Abdominis on Leaning Backwards and Forwards during Standing in Man. (From an Experiment by W. F. Floyd and P. H. S. Silver.)

Muscle action potentials recorded with surface electrodes.
At the beginning of the record the man is standing comfortably. Note the absence of muscle action potentials in erector spinæ and rectus abdominis.
The deflections marked X are the RS waves of the electrocardiogram and should be ignored.
At first signal, *lean backwards*. Note :
(i) Initial brief activity of erector spinæ which pulls the trunk backwards ; the muscle then relaxes again.
(ii) This is followed by sustained activity of the rectus abdominis which maintains the posture against gravity.
At second signal, *lean forwards*. Note :
(i) Momentary increase in activity of rectus abdominis which pulls the trunk forwards ; the muscle then relaxes.
(ii) This is followed by sustained activity of the erector spinæ which maintains the posture against gravity.
[When the subject bends forwards sufficiently to touch his toes the activity of the erector spinæ suddenly ceases ; the posture then has to be maintained purely by the tension of the ligaments of the spine.]

rectus abdominis first contracts; gravity then takes over and would make the subject fall forwards. The erector spinae contracts either to hold the forward-bent position or to enable the forward movement to proceed smoothly.[1]

(*b*) Abduct the left arm ; this causes the trunk to tend to fall to the left. The antagonist to this movement is the opposite (right) external oblique which contracts to restore and maintain the erect position (Fig. 423). When

[1] " When a person bends forward from the erect position without having to overcome an obstacle it is not the flexors of the spine which act but the extensors ; and conversely in extending the spine in leaning backwards from the erect position it is not the extensors of the spine which act but the flexors.

" For every slow, unresisted movement which is made in the direction of gravity the muscles which act in the direction of gravity are relaxed while their antagonists contract and support the part." (Beevor.)

both arms are abducted simultaneously, the trunk remains well-balanced and no contraction of the external obliques occurs.

(3) Synergic or Fixing Muscles (Synergists).—In clenching the fist, the synergic muscles are the extensors of the wrist which greatly facilitate the movement. Anyone can readily demonstrate for himself that when the fist is normally clenched the wrist becomes dorsiflexed and that fist clenching becomes uncomfortable if the wrist is held quite horizontal. In very firm closure of the fist, the triceps contracts too, acts as a synergist, and so

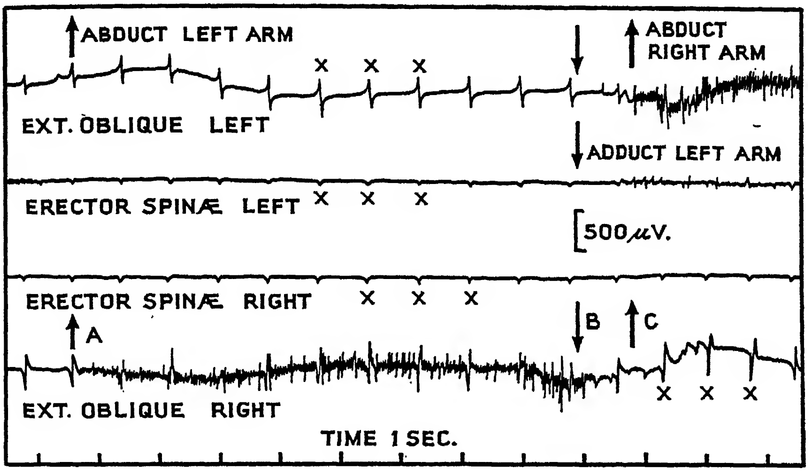


Fig. 423.—Pattern of Muscle Activity of Trunk Muscles on Abducting Arm. (From an Experiment by W. F. Floyd and P. H. S. Silver.)

Muscle action potentials recorded with surface electrodes.
The man is standing comfortably. Note absence of activity in trunk muscles.
The deflections marked X are the RS waves of the electrocardiogram and should be ignored.
At first arrow, A, the left arm is abducted into the horizontal position; the centre of gravity is thus shifted to the left. The muscle which opposes the action of gravity is the opposite external (right) oblique. This antigravity muscle contracts to maintain the erect posture.
At second arrow, B, adduct left arm. Right external oblique relaxes.
At third arrow, C, abduct right arm. Opposite (left) external oblique contracts.

prevents joints other than those actually engaged in the desired movement from being moved (in this case the elbow-joint).[1]

[1] It is desirable to note here that certain groups of muscles are normally associated together in carrying out voluntary movements (Wilson):

(i) *Head, Face, and Neck.*—(*a*) Shut the eyes: the eyeballs roll up. (*b*) Turn the eyes to the side: the head turns also. (*c*) Open the mouth against resistance: the head extends. (*d*) Turn the eyes up: the forehead wrinkles, the head extends. (*e*) In the supine position: flex the head on the chest, the recti abdominis contract.

(ii) *Arm.*—(*a*) Abduct the extended arm at the shoulder: the opposite external oblique contracts (Fig. 423). (*b*) Open the closed hand: the flexors of the wrist contract, the extensors of the wrist relax. (*c*) Abduct the little finger: the abductor ossis metacarpi pollicis contracts.

(iii) *Trunk and Legs.*—(*a*) Supine position: lift up one extended leg (*i.e.* flex the limb at the hip): the contralateral hamstrings contract. (*b*) Erect position: abduct one extended leg at the hip: the opposite tensor fasciæ femoris contracts.

(4) Movements of Cooperation.—These may be defined as movements which precede or accompany the prime movement and make it easier to carry out.[1]

It is clear from this discussion that a "volitional" movement involves the protagonists, the coordinated action of the antagonists, the assistance of the synergists to fix certain joints, and the help of various extraneous groups of muscles to carry out movements preceding or accompanying the prime movement in order to make the latter as easy and effective as possible. The component parts of the movement must also follow one another in appropriate sequence, and the whole movement must be correctly related to the stimuli which arouse it. Further, the movement is superimposed on a *basis of efficiently carried out postural activity.*

Disturbances of Voluntary Movement.—Voluntary movements may be disturbed by *abnormal function of those parts of the cerebral cortex which lie between the sensory receptor areas and the executive precentral cortex.*

(1) If the nature of the object is not recognized—the condition called *agnosia* is present. Thus a man shown a pencil does not recognize it as such, but thinks it is a cigar and uses it accordingly. The movement is planned and executed correctly, but the effect is ludicrous, because the idea in the patient's mind of the nature of the object was incorrect.

(2) The nature of the object may be recognized, but the main idea of the movement cannot be analysed into its component parts. A man was supplied with a cigar and matchbox, both of which he recognized. When asked to use them, he opened the matchbox and stuck the cigar in and tried to shut the box as though it were a cigar-cutter. Then, taking the cigar out, he rubbed it on the side of the box as though it were a match. The main idea of the action was to light the cigar, but the components of the idea were not directed to the proper object nor followed in the right order.

(3) Lastly, the nature of the object may be recognized, the general idea of action is normal, but the appropriate impulse patterns do not reach the precentral cortex intact—a normally worked-out idea is not exteriorized into the corresponding pattern of movements. Such a condition is termed *motor apraxia.* The nature of the disturbance is best illustrated by quoting Liepmann's classical case. The case was one of right-sided motor apraxia, the left side being normal. There was no paralysis on the right side. When asked to brush the examiner's coat, he took the lower corner of the coat correctly with the left hand and held it; but with the right hand he picked up the brush and made a rhythmical series of movements in the air above his right ear. At the telephone, he put the receiver to his ear with the left

[1] The following examples may be given: (*a*) Order a person sitting in a chair to stand up. The essential movement is to rise from the chair. Before that is attempted, however, he brings his feet in underneath the body to a point almost under his centre of gravity; this in normal people forms an essential preliminary to rising. Again, a hand is often stretched out to grasp the side of the chair to help the body up.

(*b*) Attempt to turn round in the chair with the hands folded on the chest. If the movement is to the right, the subject first abducts his right leg; as this "auxiliary" movement is closing, he carries out the prime movement of turning the body. As this is being completed, he adducts the left leg up to its fellow.

(*c*) Arm-swinging in walking is a movement of cooperation helping to maintain the balance of the body. It can be voluntarily dispensed with, and one can walk with the arms behind the back, but there is consciousness of a sense of lack of ease. An athlete uses this movement of arm-swinging in a powerful manner to help propel himself along.

hand, but with the right hand he took the mouthpiece and put it to his forehead, making nodding and puffing movements all the while. Another patient who showed similar difficulties said, " Je comprends bien ce que vous voulez, mais je ne parviens pas à le faire."

It has been suggested that an area in *the left prefrontal lobe* controls skilled movements via the motor cortex of *both* sides of the brain. The left motor cortex would be reached by short connecting fibres, while the fibres for the control of the right motor cortex would pass through the *corpus callosum.* Motor apraxia has been found following lesions of the left prefrontal cortex or subcortex and of the corpus callosum.

Speech.—Spoken and written speech are important examples of skilled and elaborate voluntary movements. As with other voluntary movements, all forms of speech depend primarily on the integrity of the afferent mechanisms. For spoken speech we must first be able to hear sounds—that necessitates an intact auditory pathway from the ears to the auditory centres ; secondly, we must be able to understand them—this process is believed to be related to the activity of the adjacent " auditory-psychic " areas ; the term " auditory speech centre " is used to describe the special cortical region concerned. Similarly, for written speech we need an intact pathway from the eyes to the visual cortex to enable us to see ; the written symbols must be correctly interpreted. This process is believed to be related to the activity of the " visual speech centre " (which is presumably part of the general " visuo-psychic " area). The sensory speech centres are mainly found on the *left* side of the brain in right-handed individuals. The exteriorization of speech demands the skilled use of many muscles, such as those of the tongue, larynx, or hand. The excitomotor areas of the brain, their descending tracts, and the related lower motor neurones, must all be intact and functioning normally, and appropriately guided by the cerebellum and the muscular, labyrinthine, and other afferents. But it is further supposed that, as for other voluntary movements, there is need for some higher centre which mediates the " planning " of the details, sequence, and duration of the various movement patterns employed. It is suggested that there is a " motor speech centre " for both spoken and written speech in the left prefrontal lobe, in the neighbourhood of the alleged higher centre for voluntary movements in general. The " higher " motor centre for spoken speech is often called Broca's centre, after the man who first attempted to locate it with some precision. He suggested that it lay in the left inferior frontal convolution.[1]

The sensory speech centres are undoubtedly linked by important pathways with the motor centres and guide them in their activities.

[1] Stimulation of the cerebral cortex in conscious man never produces spoken speech ; this result is not surprising as cortical stimulation only produces " fragments " of even quite simple movement patterns. Stimulation of the lower part of the motor face area (and the posterior part of the superior frontal convolution) may give rise to a " cry " (like that preceding an epileptic fit) ; the cry is called " vocalization " but it is not speech. Excisions in man in the vicinity of area 44 produce motor aphasia of varying degrees of severity and duration ; Penfield suggests that this region is probably Broca's area. Stimulation of this area in conscious man disturbs its function and produces arrest of speech ; one patient said later, " I knew what I wanted to say but could not " (cf. the comment of the patient with apraxia (*supra*) " je ne parviens pas à le faire "). Stimulation of the posterior part of the parietal lobe or of the upper temporal lobe also produces " aphasic arrest " (sudden speechlessness) suggesting that these regions are also involved in speech.

The above schema can be summarized as follows.

a'

EARS——→Auditory Centres—|→" Auditory Speech Centre "

—→" Motor Speech Centres"

EYES——→Visual Centres—|→" Visual Speech Centre "

a

—|→Excitomotor Areas——→Lower Motor Neurones——→Muscles.

b

This schema is too rigid, and underestimates the closeness of co-ordination that exists between various regions of the cerebral cortex. It is probably unjustified to define and locate speech centres with any precision or to assign to them relatively autonomous functions. The *left side of the brain* is far more important than the right in *right-handed individuals* for the carrying out of speech functions in general. Marie stressed that speech functions must be considered as a *whole,* both on the receptive and the expressive side, and both for spoken and written speech. It is best to talk of a " cortical speech area " or " region " in a vague way rather than to use the terms " speech centres," which give a misleading idea of non-existent precision of localization. As speech is dependent on visual, auditory, and proprioceptive impulses, the speech area must be in the closest functional and anatomical connections with those regions of the cortex which primarily receive these impulses. The temporo-sphenoidal lobe, the hinder part of the parietal lobe (including the supra-marginal and angular gyri), and the island of Reil, probably represent the principal receptive side of the speech area. The area is continued forward to include that part of the prefrontal lobe from which the efferent pathway may possibly emerge. The whole of this vast area is correlated with the highest intellectual activities ; and injury to this region depresses not only speech but also other intellectual activities. Injuries to the *incoming* pathways (auditory, visual, or proprioceptive), to the *outgoing* pathway to the excitomotor areas, or to the *cerebral grey matter* itself, cause fairly characteristic disturbances of speech, which are discussed more fully below. In general, it may be said that injury to the path to the visual speech centre (*a* in schema) may produce *word blindness* (inability to understand written words). Injury to the outgoing pathway to the excitomotor areas (*b* in schema) may produce *pure motor aphasia* (inability to speak articulately), or inability to write (*agraphia*), or both, without much mental disturbance. Otherwise most cases of aphasia which result from injuries to the cerebral grey matter itself, usually show interference with both the expressive and the appreciative side of speech and deterioration of the intellectual faculties. Thus, simple sentences may be

understood, while those involving more complex ideas cannot be followed. The patient may obey when told to raise his hand, but does not know what to do if he is asked to touch his right hand with his left forefinger.

Clinical Conditions.—PURE MOTOR APHASIA is a condition in which there is loss of articulate speech without mental confusion or deterioration. The patient is dumb, though the excitomotor cortex and its efferent paths are intact. The condition can be regarded as a special example of *apraxia* (p. 652); a certain highly complicated group of movements, *i.e.* those of speech, are lost. It is commonly due to a lesion, *subcortical* in position, which cuts off the speech area from the excitomotor cortex.

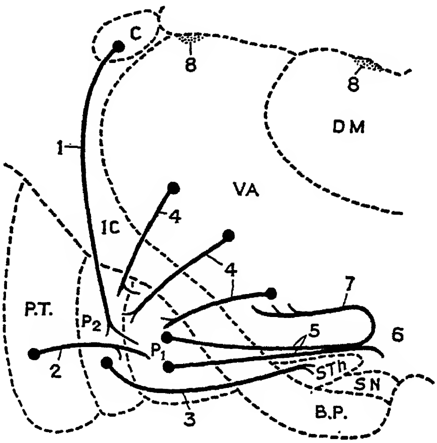


FIG. 424.—Connections of Corpus Striatum. (Ranson, *Res. Publ. Ass. nerv. ment. Dis.*, 1941, *21*, 737.)

B.P., basis pedunculi.
C, caudate nucleus.
DM, dorsomedial nucleus of thalamus.
IC, internal capsule.
P.T., putamen.
P_2, P_1, external and internal division of globus pallidus.
SN, substantia nigra.
STh, subthalamic nucleus.
VA, anteroventral nucleus of thalamus.
1, fibres from caudate nucleus to globus pallidus [striopallidal].
2, fibres from putamen to globus pallidus [striopallidal].
3, fibres from globus pallidus to subthalamic nucleus [pallido-subthalamic].
4, fibres from thalamus to globus pallidus [thalamopallidal].
5, fasciculus lenticularis.
6, fibres from globus pallidus to hypothalamus [pallido-hypothalamic].
7, fasciculus thalamicus.
8, pallido-habenular tract.

PURE WORD-BLINDNESS is the inability to recognize the meaning of written or printed words, which appear as hieroglyphics. The patient is unable to read aloud or copy print into writing. Speech otherwise is not disturbed. Word-blindness (*alexia*) is a special form of *visual agnosia*; there is a loss of the ability to recognize familiar symbols. It is due to a lesion, subcortical in position, which cuts off the visual centres from the speech area. It is usually situated deep in the substance of the occipital lobe behind the angular gyrus, and often involves the optic radiation, which is close by and is passing to the occipital cortex. Word-blindness is thus often associated with right hemianopia (cf. p. 580).

PURE WORD-DEAFNESS, *i.e.* the failure to recognize the meaning of spoken speech with no other defect of speech or intelligence, does not occur clinically (see Syntactical Aphasia, p. 656, footnote 1, (3)).

CORTICAL APHASIA.—As already explained, pure word-blindness and pure motor aphasia are caused by *subcortical* lesions. Injuries to the *cortex* depress speech function as a whole. Lesions situated posteriorly affect more markedly the *reception* of speech, while those situated anteriorly disturb, in the main, the *execution* of speech. The later acquired speech functions tend to drop out before those developed in early life. Propositions cannot be

made when ejaculations and emotional expression are still possible; the latter being more deeply grounded, persist longer.[1]

Dysarthria.—Dysarthria is a disturbance of speech resulting from defects of the *excitomotor areas in the cortex and their connections*, *i.e.* the pyramidal tracts, cranial nuclei, cranial nerves, muscles; from disturbances of tone, like rigidity; or from lack of cerebellar control of postural and voluntary activity. Any of these disorders or a combination of them disturb speech to some extent; it should be noted that the muscles cannot be used effectively for *any* purpose, including that of speech.[2]

THE CORPUS STRIATUM

The corpus striatum consists of the *caudate* and *lenticular* nuclei. The caudate nucleus is part of the cerebral suppressor or inhibitory mechanisms; the lenticular nucleus is part of the cerebral excitatory or *facilitatory* mechanisms. The connections of the corpus striatum are shown in Figs. 424, 425.

(1) CAUDATE NUCLEUS.—(i) As explained on p. 620, the caudate nucleus lies on the path by which the cortical suppressor bands inhibit the thalamus and so abolish the cortical resting potentials.

(ii) Inhibitory descending fibres from the suppressor bands relay in the caudate nucleus to pass on to the reticular nuclei and thence via the reticulospinal tracts to the lower motor neurones (Fig. 400; p. 623).

(iii) The caudate nucleus also receives afferents from the thalamus; it sends efferents to the globus pallidus [striopallidal tract].

(2) LENTICULAR NUCLEUS.—This is subdivided into an external segment the *putamen* and a smaller internal part, the *globus pallidus*.

(i) The extrapyramidal fibres arising in areas 4 and 6 relay in the putamen and globus pallidus to pass to the *hypothalamus* and the excitatory *reticular*

[1] Head recognizes four clinical varieties of aphasia

(1) VERBAL APHASIA.—There is defective word formation; words both in internal and external speech are evoked with difficulty. Enunciation is slow and halting; writing is difficult or almost impossible. Numerous errors are present in articulate speech. This corresponds roughly to motor aphasia and agraphia (*b* in schema).

(2) NOMINAL APHASIA.—There is difficulty in naming letters, words, or objects. Reading is impaired. This occurs especially with lesions of the upper temporal lobe.

(3) SYNTACTICAL (OR "JARGON") APHASIA.—When we speak aloud we immediately recognize what has been said, and are able to correct any faults made. This "backlash" keeps speech at a high functional level. Similarly, impulses from the functioning muscles keep us informed as to how the movements are executed, and visual appreciation maintains a check on writing. When the speech area is isolated on the afferent side, a voluble jargon results. The patient neither hears nor feels that he has spoken incorrectly; in these circumstances speech is greatly deranged. The patient becomes excited and very voluble, and speaks in a gibberish from which distinguishable words may be entirely absent. He is usually deemed mad by those around him (Collier).

(4) SEMANTIC APHASIA (semantic: concerned with "meaning").—There is want of recognition of the full significance of words and phrases.

[2] The clinical causes of dysarthria may thus be summarized:

(1) *Muscles:* paralysis of muscles of face in myopathy or myasthenia gravis.
(2) *Peripheral nerves:* bilateral facial palsy; diphtheritic neuritis.
(3) *Cranial nuclei:* bulbar form of progressive muscular atrophy; tumours of medulla.
(4) *Pyramidal system:* *e.g.* spastic bulbar paralysis; general paralysis.
(5) *Cerebellum:* *e.g.* disseminated sclerosis, Friedreich's disease.
(6) *Extrapyramidal system:* from rigidity in Parkinson's disease.

nuclei; reticulospinal fibres descend to reach the lower motor neurones which they stimulate (Fig. 409; p. 629).

(ii) The other connections are:

(a) Two-way connections between the globus pallidus and the thalamus (mainly anteroventral nucleus).

(b) From the globus pallidus to the hypothalamus and subthalamic nucleus.

(c) To the substantia nigra.

(d) From the putamen to the globus pallidus [striopallidal tract].

Parkinson's Disease [Paralysis Agitans].—In this disorder the subcortical nuclei, the substantia nigra, and globus pallidus are most commonly damaged. A study of the clinical findings may throw some light on the functions of the corpus striatum and the related subcortical nuclei. The characteristic signs of Parkinson's disease are: rigidity; involuntary movement—tremor; disturbances of movement.

(1) RIGIDITY (STRIATAL RIGIDITY).—The rigidity (the clinical term for the *increased muscle tone* which is present) affects mainly the big muscle masses and the proximal parts of the limbs. It affects both protagonists and antagonists but as the rigidity is usually greater in certain groups of muscles, a characteristic posture results. Thus, the sternomastoids, the biceps, and knee flexors are often markedly involved; the back is slightly flexed, the arms are adducted and flexed and the knees are bent, *i.e.* an attitude of generalized flexion is present. The rigidity, like all patterns of skeletal muscle tone is due to proprioceptive reflexes; if procaine is injected into the nerve supply of a muscle in a concentration which paralyses only the muscle afferents, the rigidity is abolished. The hypertonus of striatal rigidity differs in several ways from the "spasticity" which accompanies hemiplegia: (i) the distribution is different (cf. p. 642); (ii) if the limb is progressively flexed or extended, resistance is encountered all the time, and at no stage are the muscles reflexly inhibited, *i.e.* there is no lengthening reaction; (iii) the rigidity is unaffected by neck reflexes, *i.e.* by moving the head on the trunk (cf. p. 591). Striatal rigidity, like spasticity, must be regarded as a release phenomenon. It is not clear what exactly are the higher level inhibitory mechanisms which the disease has put out of action; the caudate nucleus is however part of the descending pathway inhibiting muscle tone (p. 623).

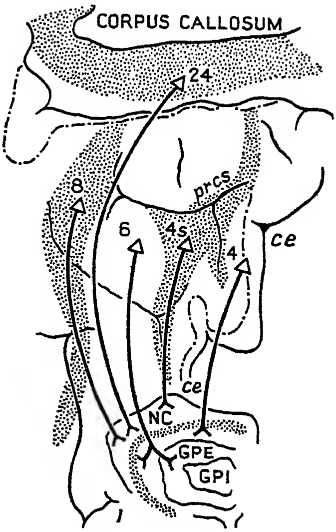


FIG. 425.—Connections between Precentral and other parts of the Cerebral Cortex and the Corpus Striatum (chimpanzee). (McCulloch, in Bucy, *Precentral Motor Cortex*, 1944).

Shaded areas (24, 8, 4s) are supressor bands. Note that the supressor bands connect with and "fire" the caudate nucleus (NC). The excitomotor areas 4 and 6 connect with and "fire" the globus pallidus (GPE, GPI) and the putamen (lying between NC and GPE) of the lenticular nucleus.

The rigidity in advanced cases may be so marked that the patient, if not too heavy, can be carried about like a statue. Any grade of rigidity

must obviously interfere with the ease and rapidity of execution of movements; intense rigidity may fix the joints and make voluntary movement impossible.

(2) Tremor.—The tremor consists of fairly regular, rhythmically alternating contractions of a muscular group or groups and their antagonists, usually about six times per second. The range varies from a very fine to a coarse, wide movement. Though any group of muscles and their opposites may be affected the tremor commonly involves the more distal and mobile segments of the body, such as the fingers, hand, lips, or tongue. It may consist of a flexor-extensor movement at the wrist, a side-to-side movement (radio-ulnar deviation), or a movement of pronation-supination. The tremor is readily modified by many conditions. It may be temporarily inhibited by painful stimuli; it ceases during sleep. The tremor must presumably be regarded as a release phenomenon but its mode of production is unknown.[1] (Experimentally produced tremor of a similar type is considered on p. 660.)

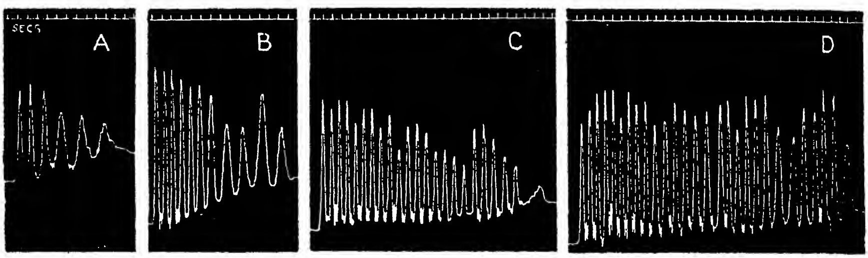


Fig. 426.—Rigidity interfering with Movement in Parkinson's Disease.

Series of records of voluntary flexion-extension of right elbow. A. Beforei njection. B. Ten minutes after injection of 20 c.c. of 1% procaine solution into biceps to paralyse *afferent* fibres in motor nerve. C. Fifteen minutes after injection into biceps. D. Twenty-five minutes after injection into biceps and six after injection into triceps. In A, movements dwindle almost immediately in speed and range; there is considerable improvement in B and C; after paralysing afferents in triceps (D), movements are rapidly performed and amplitude is sustained. (Walshe, *Lancet*, 1920.)

(3) Disturbance of Movement.—The disturbances of movement in Parkinson's disease are as follows:

(i) There is *weakness* of the muscles; *fatigue* sets in rapidly; there is difficulty in maintaining a contraction; the movements are *slow*, irregular, and limited in extent. This is well seen in the case of the interossei: on attempting to spread out the fingers the movement is not readily carried out, and the feebleness and lack of precision are well marked. These defects involve chiefly the *small* muscles, or at any rate are less evident in the larger masses of muscle. The voice is monotonous, and the "chanson de langage" is lost. The delicate adjustments of the laryngeal muscles necessary to produce the inflections and modulations of normal speech are lost. There is jerky movement of the eyes on looking to one side. It is thus seen that muscles are affected chiefly in inverse proportion to their mass. All the small muscles are affected whether they are used in highly complex movements (*e.g.* larynx) or in relatively simple movements (*e.g.* interossei). The

[1] Tremor also occurs clinically at rest (static tremor) and during movement ("intention" tremor) in cerebellar dysfunction (pp. 609 and 610), in old people, and in exophthalmic goitre (p. 991).

impairment of voluntary movement is greatly influenced by the degree of rigidity. When an antagonistic pair of muscles (*e.g.* biceps and triceps) are deafferented by means of procaine, the *motor* fibres being left intact, the rigidity is abolished and it is found that the normal range and speed of movement are temporarily restored (Fig. 426).

(ii) *Poverty of Movement.*—When at rest the patient with Parkinson's disease literally sits still and does not make the spontaneous movements usually shown by normal people. A man does not cross his legs or toy with his lips and chin during conversation ; a woman does not twirl her wedding-ring. Intelligent patients explain they have no desire to move. The rigidity when present is a contributory factor in making movement difficult and irksome, but it is not an essential accompaniment of the poverty of movement. *Movements of cooperation* (p. 652), *e.g.* arm-swinging, are similarly, less in evidence.

(iii) *Emotional Expression.*—The face is *mask-like* in appearance, though the patient *can* and does laugh or cry. Emotional movements are carried out slowly and less effectively. One patient said, " I used to have a ringing laugh, now I only make a silly noise." There is *not an absence* of expression but a lack of *ease* (due partly to rigidity) in passing from one expression to another.

(iv) *Reflexes.*—The reflexes, both superficial and deep, are unaffected at first. When the rigidity becomes severe in degree the joints become fixed, and this mechanically interferes with the reflexes. The ultimate clinical picture in progressive cases is one of complete motor helplessness. The rigidity increases and binds down the joints. Speech and swallowing become difficult, and finally impossible. Tremor is usually marked.

Experimental Parkinsonism.—Attempts to reproduce the clinical syndrome in animals have given the following results :

(1) Lesions of the *caudate* nucleus or the *putamen* do not give rise to Parkinsonism. There are conflicting reports about the effects of bilateral lesions of the *globus pallidus* : some workers (Ranson) have obtained no characteristic changes ; others have noted striking poverty of movement, without paralysis or rigidity.

(2) (i) *Bilateral* lesions (in the *monkey*) of the *hypothalamus* involving the lateral areas where, incidentally, the fibres from the globus pallidus to the hypothalamus lie, produce interesting results. There is again no paresis, no tremor, chorea, or athetosis, and no rigidity. Although the animal can run, jump, or climb, its movements are slow and there is poverty of movement. The animal stays quite still for long periods without change of position ; apart from opening and closing the eyes, no movement may be observed for minutes on end. The face is described as displaying a " fixed sad expression which does not change." The *grasp reflex* is released : if a rod is placed in the hand, it is grasped so firmly that the animal can be lifted up by means of the rod.

(ii) *Bilateral* lesions (in the *cat*) in a position close to that discussed above, *i.e.* between the caudal border of the mammillary bodies and the rostral fibres of the third nerve, produce a somewhat different picture. The animal is very lethargic and shows " catalepsy " ; it remains in any position in which it is placed. With the animal lying on its back in a shallow trough, if the limbs are flexed they remain flexed, if extended they remain extended ;

there is no tendency (as in decerebrate rigidity) for the limbs to return to a basic posture. Most of the time the animal seems to be asleep; on pinching the tail, it wakes up, assumes a normal posture, takes a few steps, and then often goes to sleep again in a standing position. When awake there is great paucity of voluntary movement.

It was pointed out on p. 630 that the globus pallidus and hypothalamus form part of a descending excitatory pathway which facilitates the response of the lower motor neurones to impulses in the pyramidal tracts. The experiments described above suggest that interruption of this pathway may be an important factor in producing the weakness and paucity of movement of Parkinson's disease. It could not, however, account for the characteristic rigidity which must be due to loss of some tonic inhibitory influences.

(3) TREMOR.—An alternating tremor resembling that seen in clinical Parkinsonism has been produced experimentally by means of lesions in the tegmentum of the midbrain and pons (but not by lesions involving the basal ganglia). The rhythm of the tremor is not visible in the electro-encephalogram experimentally or clinically, indicating that it is *not* due to an alternating discharge from the motor cortex; it is suggested that the tremor is due to some abnormal pattern of activity of reticulospinal excitatory and inhibitory pathways.

It must be frankly admitted that our knowledge of the corpus striatum and the related nuclei is too fragmentary to enable us to give any coordinated account of their functions or to account adequately for the signs in Parkinson's disease in man.

Clinical Conditions.—The corpus striatum is probably involved in the following diseases:

(1) PARKINSON'S DISEASE (paralysis agitans), which may be the result of vascular lesions of this region occurring in old people.

(2) ENCEPHALITIS LETHARGICA: the corpus striatum is very commonly affected, producing an end result almost indistinguishable from the previous condition, except that it comes on acutely in younger subjects.

(3) PROGRESSIVE LENTICULAR DEGENERATION (Wilson), a bilateral affection of the lenticular nuclei associated with multilobular cirrhosis of the liver.

(4) KERNICTERUS:[1] the basal ganglia and related parts of the brain are damaged and stained yellow in hæmolytic disease (p. 180), which results from the noxious action of Rhesus antibodies. The nuclei usually involved are the globus pallidus and the subthalamic body (the putamen is affected more rarely). If the child survives it may show the striatal type of rigidity, chorea, and athetosis (and often mental deficiency).

Chorea and Athetosis.—These two disorders are characterized by spontaneous movements; clinically lesions are usually found in the thalamus or corpus striatum.

In *chorea* the spontaneous movements are irregular, brief, and rapid. If the arm is affected there is, in addition, decreased tone and muscular weakness, especially of the hand (there is complaint of "uselessness" of the hand); the arm hangs limply. Voluntary movements of the arm are not carried out smoothly but tend to be abrupt and sudden.

In *athetosis* the spontaneous movements are slow and confluent—one

[1] Evans and Polani, *Quart. J. Med.*, 1950, *19*, 129.

phase blending with the next; the limb thus undergoes a series of complex writhings and contortions. As Wilson puts it, the movements appear to be directed to an end which is never attained.

The pattern of the involuntary movements in chorea and athetosis suggests that they arise in the cerebral cortex; they only occur when there is relative integrity of the descending paths from the excitomotor cortex.

Experimentally it is claimed that an isolated lesion of the caudate nucleus produces choreo-athetosis. On the basis of these results it has been suggested that the spontaneous movements are due to interruption of the suppressor pathway from area 4s via the caudate nucleus to the thalamus, thus blocking the thalamocortical circuit and depriving the excitomotor area of an afferent control from the subcortical ganglia. Wilson has suggested that a lesion of the thalamus might deprive the cerebral cortex of cerebellar control or lead to disordered cerebellocortical influences and thus give rise to the spontaneous movements; he draws attention to the other signs suggestive of cerebellar dysfunction which are found in chorea, *e.g.* decreased muscle tone, weakness and incoordination of willed movements.

It is claimed that the movements of athetosis may be controlled by (i) removal of area 6, or (ii) section of the ventral column of the spinal cord, without encroaching on the crossed pyramidal tracts.[1]

It is clear from the discussion that it is premature to assign a precise pathology to chorea and athetosis.

PHYSIOLOGY OF THE EMOTIONS

"Of points where physiology and psychology touch, the place of one lies at 'emotion'." A discussion of the physiology of the emotions must begin by recognizing that man is body and mind, a "psycho-biological whole," and that emotion has a mental and a physical side. The mental side consists of *cognitive, affective,* and *conative* changes.[2] The following example illustrates the meaning of these terms: thus, I hear a noise which I recognize as that of an exploding bomb (cognition); I feel frightened (affect), and I want to take shelter (conation). The physical side of an emotion consists of changes in viscera and skeletal muscles; these are often widespread and involve the coordinated activity of both the autonomic and somatic nervous systems.

Bodily Changes in Emotions.—The bodily changes accompanying several "coarse" emotions are described below.

FEAR.—"In fear the eyes and mouth are widely opened and the eyebrows raised. The heart beats quickly and violently, so that it palpitates or knocks against the ribs; the skin becomes pale owing to the vasomotor centre being affected in such a manner as to cause contraction of the small arteries of the skin; perspiration immediately exudes from the skin, which is cold, hence the term, a cold sweat. The hairs on the skin stand erect, and the superficial muscles shiver. The breathing is hurried. The salivary glands act imperfectly; the mouth becomes dry and is often opened and shut; under slight fear there

[1] Bucy, *J. Neurol, Neurosurg, Psychiat.*, 1951, *14*, 108.

[2] Cognitive, affective, and conative are classical psychological terms. Cognition is derived from the Latin "cognoscere" (to get to know); an "affect" "pertains to the emotions" (cf. "affectionate"), it is a "feeling"; conation is the faculty of volition and is related to desire, urge, drive.

is a tendency to yawn. One of the best-marked symptoms is the trembling of all the muscles of the body; and this is often first seen in the lips. From this cause, and from the dryness of the mouth, the voice becomes husky or indistinct or may fail altogether. As fear increases into an agony of terror, we behold, as under all violent emotion, diversified results. The heart beats wildly or may fail to act and faintness ensues; there is a death-like pallor; the breathing is laboured; the wings of the nostrils are widely dilated; there is a gasping and convulsive motion of the lips, a tremor on the hollow cheek, a gulping and catching of the throat; the uncovered and protruding eyeballs are fixed on the object of terror, or they may roll restlessly from side to side. The pupils are said to be enormously dilated. The arms may be protruded as if to avert some dreadful danger, or may be thrown wildly over the head." (Darwin).

Grief.—" The chief feature in the physiognomy of grief is its effect on voluntary movements. There is a feeling of weariness; movements are made slowly, without strength, and with exertion, and are limited to the fewest possible. By this the grieving person gets his outward stamp; he walks slowly, unsteadily, dragging his feet and hanging his arms. His voice is weak and without resonance; he prefers to sit still, sunk in himself and silent. The tone of the muscles is diminished. The neck is bent, the head hangs (' bowed down ' with grief). With this condition of weakness of the muscles there coexists a subjective feeling of weariness and heaviness, of something which weighs upon one; one feels ' downcast,' ' oppressed,' ' laden,' one speaks of his ' weight of sorrow.' Many ' succumb ' to sorrow to such a degree that they cannot stand upright but sink or lean against surrounding objects, fall on their knees, or, like Romeo in the monk's cell, throw themselves upon the earth in their despair. But this weakness of the entire voluntary motor apparatus is only one side of the physiology of grief. Another side, belongs to the involuntary muscles, especially those in the walls of the blood vessels. The vascular muscles are strongly contracted so that the tissues and organs of the body become anæmic. The immediate consequence of this bloodlessness is pallor and shrunkenness, and the pale colour and collapsed features are the peculiarities which, in connection with the relaxation of the visage, give to the victim of grief his characteristic physiognomy. Another regular consequence of the bloodlessness of the skin is a feeling of cold and shivering. Some secretions are diminished; the mouth grows dry, the tongue becomes sticky. Other secretions are increased; there is the weeping, with its profuse secretion of tears, its swollen reddened face, red eyes, and augmented secretion from the nasal mucous membrane." (Lange).

Rôle of Sympathetic and Parasympathetic Systems in Emotional Reactions.—(1) As already explained, in some emotional states, *e.g.* in fear, the *sympathetic* nervous system is stimulated as shown by a quick heart, vasoconstriction and rise of blood pressure, dilatation of the pupil and increased secretion of adrenaline.

(2) In many emotions, however, *parasympathetic* overactivity predominates. (i) The anxious candidate awaiting a viva-voce examination suffers from frequency of micturition; the sacral autonomic is motor to the bladder. (ii) Victorian young ladies used to swoon on the slightest provocation; the swoon is the vaso-vagal syndrome (p. 271) in which the heart is slowed by vagal overactivity and the vasomotor centre is depressed, with a consequent

decreased discharge of sympathetic vasoconstrictor impulses to the blood vessels. (iii) The vagal innervation of the stomach is stimulated in states of anxiety and resentment, leading to increased secretion of gastric juice and dilatation of the blood vessels of the mucosa (p. 520).

(3) In some emotions there may be *complex* autonomic activity, *e.g.* increased sympathetic action in one viscus and increased parasympathetic action elsewhere. Thus in grief the sympathetic nerves constrict the skin blood vessels while the parasympathetic nerves are responsible for the secretion of tears and the dilatation of the blood vessels of the eyes. In coitus the sacral autonomic is responsible for pelvic vasodilatation and erection of the penis while the sympathetic produces cardiac acceleration and a rise of blood pressure.

It is clear that in emotional states the *appropriate* parts of the autonomic system, sympathetic as well as parasympathetic, are employed to produce the characteristic and presumably appropriate visceral reactions.

INFLUENCES OF EMOTIONAL STATES ON DUCTLESS GLANDS.—Emotional states influence the secretion of adrenaline (p. 731) and of the antidiuretic hormone of the posterior pituitary (p. 53). There is evidence that the brain regulates the discharge of the anterior pituitary hormones (p. 931); emotional states might in this way indirectly affect the activity of the thyroid (*e.g.* possibly producing hyperthyroidism (p. 980)) and adrenal cortex, and the secretion of the hormones of the ovary and the testis.

Central Control of Emotional Exteriorization.—Certain regions of the cerebral cortex (especially the precentral and prefrontal areas), the thalamus, hypothalamus, and brain stem are concerned in regulating and mediating the complex coordinated patterns of bodily activity which are characteristic of emotional states.

1. **Rôle of Cerebral Cortex.**—(1) Evidence set out on p. 670 shows that stimulation or excision of parts of the prefrontal lobes (especially areas 13 and 14 on the orbital surface or area 24 [cingular gyrus]), may enhance or depress the outward manifestations of emotion, alter patterns of behaviour and produce changes in the personality.

(2) Following *lesions of the pyramidal tracts*, especially in cases of bilateral lesions of the corticobulbar fibres which control the motor cranial nuclei, skeletal muscles which are no longer responsive to voluntary control " spring into action and give free expression to sadness or gaiety." In a typical case there is loss of voluntary movements of the face, but emotional expression is preserved intact and is elicited in an exaggerated form by what are normally inadequate and inappropriate stimuli. The patient may thus display well-executed, prolonged fits of laughing, crying, scowling, or frowning which he cannot check or control. One patient started laughing in this way (for no particular reason) at 10 a.m. and continued with few pauses till after midnight.[1] It is important to emphasize that the patient does not experience the feelings which his face and body so dramatically exteriorize; there has thus been a dissociation of the mental and the physical side of the emotion; the mental, especially the " affective " component of the emotion is missing. Such behaviour is often called *pseudaffective*, *i.e.* there is an outward manifestation of an emotion which is not felt (see pp. 642, 667).

[1] Compare the behaviour of an invited audience attending a variety performance given in the British Broadcasting Corporation's studios.

The clinical findings just described indicate that emotional exteriorization is essentially an *involuntary* act which can be mediated by subcortical levels; it is, however, normally modified, and to a considerable extent inhibited, by cortical influences. The extent of such cortical control varies in different peoples and different individuals; the strong silent Englishman of fiction feels deeply but keeps his emotional exteriorization under iron control; the equivalent Frenchman not only has a highly mobile, expressive face but exteriorizes his feelings by appropriate activities of most of his bodily musculature.

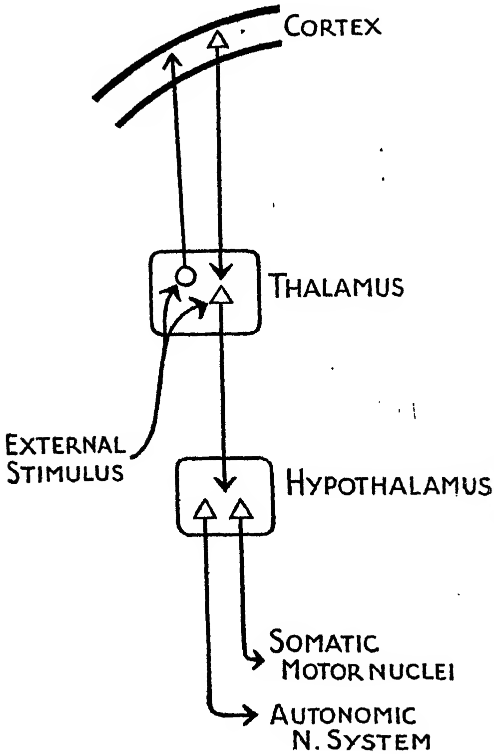


FIG. 427.—Diagram to illustrate Reflex Arcs employed in Exteriorization of Emotion.

2. **Rôle of Subcortex.**—(1) After trans-section of the brain in cats just anterior to the thalamus, there are "released" in the resulting thalamic preparation "a group of remarkable activities such as are usually associated with emotional excitement—a sort of *sham rage*." The most trivial stimuli, or no recognizable stimuli, set up widespread sympathetic activity associated with erection of the hairs, profuse sweating of the toe-pads, a rapid heart rate and raised arterial blood pressure, hyperglycæmia, adrenaline secretion and bodily movements such as lashing of the tail, kicking, running, and even biting.[1] The "width and energy" of the exteriorization "make it unmistakably the counterpart of intense fury in the normal animal." Detailed experimental ablations prove that in the cat the "rage reaction" is mediated centrally by an area comprising the caudal half of the hypothalamus and the most ventral and most caudal fractions of the thalamus.

(2) After removal of the thalamus and hypothalamus by a midbrain trans-section (midbrain or decerebrate preparation) a modified, more partial and transient rage reaction can be induced reflexly. But the response is only elicited by powerful stimulation; it consists of less widespread, *i.e.* more isolated items of behaviour; it may be limited for example to the head and show little involvement of the viscera.

(3) After a midthoracic spinal transection, pseudaffective reactions of a

[1] Compare the almost identical rage reactions displayed by the late Adolf Hitler, including the reported biting of the carpet.

still more limited character may be reflexly produced from the isolated half of the spinal cord, *e.g.* the flexor withdrawal reflex of a leg on nocuous stimulation of the sole of the foot.

It may be concluded that as more of the nervous system is available as a reflex centre, the kind of stimuli or situations which in the intact animal are associated with an " affect " (or " feeling "), elicit more complex, widespread, and varied reactions which we recognize as resembling more and more closely those outward expressions which in ourselves accompany emotions. It should be emphasized that the lower the level of a reflex " centre," the more stereotyped and restricted is the reflex response obtained. Reflexes employing higher level centres are not only more generalized but also more variable in character and seemingly better adapted to deal with the causal situation. The fact that emotional exteriorization normally varies so much in its pattern and is so nicely adjusted to the nature of the stimulus, is almost in itself adequate evidence that it is mediated by relatively high cerebral levels. The evidence presented above has made clear the rôle of the region of the thalamus and hypothalamus in mediating some of the " coarser " emotions. But even the reactions of the thalamic animal are limited in pattern, exaggerated in character, and generally ill-related to the causal agent employed. We must conclude that it is the integrity of the cerebral cortex which introduces the element of refinement, nicety of adjustment, and greater variegation of pattern into emotional exteriorization.

3. **Rôle of Hypothalamus.**—As explained on p. 716, the nuclei of the hypothalamus control both divisions of the autonomic nervous system and also influence the activity of the lower motor neurones which supply skeletal muscles. The hypothalamus is an important centre on the *efferent* (" downstream ") side of the mechanisms controlling emotional exteriorization (Fig. 427).

(1) Masserman introduced fine electrodes into the hypothalamic region of anæsthetized cats; the animals were then allowed to recover. Subsequently the hypothalamus could be stimulated at will in the conscious intact animal. Such localized electrical stimulation sets up a typical " rage reaction," both on the visceral and somatic side, indistinguishable from that observed in the thalamic animal. These responses are diminished by the administration of hypnotic drugs such as Amytal or Nembutal; on the other hand the instillation of a minute amount of Metrazol or picrotoxin initiated " a crescendo of excitement reaching a seeming frenzy of rage and fear."

(2) On one occasion de Barenne inadvertently injected strychnine into a cat's hypothalamus (intending it for the thalamus). A few moments after the cat had recovered from the ether anæsthesia " a frightening seizure of unleashed fury developed in which the cat dashed madly from one end of the room to the other and from floor to ceiling, savagely attacking anything in its path."

Central Pathways for Emotional Exteriorization.—Two-way connections are present between the prefrontal cortex and the hypothalamus (p. 670); fibres also pass both ways between the thalamic nuclei and the hypothalamus (Fig. 428). From the latter new relays pass: (i) to the appropriate supraspinal autonomic centres (*e.g.* cardiac, vasomotor); (ii) to nuclei in the midbrain and the reticular formation of the pons and medulla; from these cells, relays pass to the appropriate cranial and spinal lower motor neurones.

Fibres from the precentral and prefrontal regions of the cortex reach this region of the brain stem and modify its inherent pattern of reflex response. Lesions in man of the thalamus, certain parts of the hypothalamus, and the dorsal part of the midline region of the brain stem (reticular formation) may produce *emotional palsy, i.e. loss of emotional exteriorization, while the volitional control of the muscles used in emotional expression remains intact.*

Relation of Bodily Changes to Production of Emotional Affective State.—There are three possible relations between the emotion "felt" and the emotion "exteriorized," *i.e.* between the mental and the physiological changes:

(i) The stimulus (*i.e.* the causal situation) sets up activity of the brain which simultaneously gives rise to the feeling and the exteriorization.

(ii) The stimulus first sets up the feeling which in its turn, by influencing the brain, produces the bodily changes.

(iii) The stimulus acting via the brain first reflexly produces the bodily changes; the viscero-somatic changes thus set up send back afferent impulses which on reaching the brain induce the affective state characteristic of the emotion. This last suggestion is the James-Lange theory of the emotions. The views of James are set out more fully in the footnote.[1]

The experimental evidence summarized below suggests that James was largely in error and that he greatly overestimated the rôle of the backlash from the body to the brain in producing the "feeling" (affective) component of an emotion.

(1) As pointed out on p. 665, appropriate hypothalamic stimulation in the intact, unanæsthetized cat produces the exteriorization of rage and fear. The question arises whether the animal "felt" angry or frightened; no certain answer can be given because we cannot appeal to the cat to describe its feelings; but careful examination of the cat's behaviour during stimulation suggests that the animal is *not* feeling the emotion it is so graphically portraying. "The ostensibly aggressive activity during hypothalamic stimulation is not directed towards specific objects in the animal's environ-

[1] "The *bodily manifestations must be interposed between the stimulus and the emotion*; we feel sorry *because* we cry, angry *because* we strike, afraid *because* we tremble, and not that we cry, strike or tremble as the case may be. Without the bodily states following on the perception, the latter would be purely cognitive in form, pale, colourless, destitute of emotional warmth. We might then see the bear and judge it best to run, receive the insult and deem it right to strike, but we should not actually *feel* afraid or angry." "Objects excite bodily changes by a preorganized mechanism and the changes are so infinitely numerous and subtle that the entire organism may be called a sounding board, which every change of consciousness, however slight, may make reverberate. The various permutations and combinations of which these organic activities are susceptible make it abstractly possible that no shade of emotion, however slight, should be without a bodily reverberation as unique, when taken in its totality, as is the mental mood itself." There follows a purple passage which is an appeal to introspection: "What kind of an emotion of fear would be left if the feeling neither of quickened heart-beats nor of shallow breathing, neither of trembling lips nor of weakened limbs, neither of goose-flesh nor of visceral stirrings were present, it is quite impossible for me to think. Can one fancy the state of rage and picture no ebullition in the chest, no flushing of the face, no contraction of the muscles, no clenching of the teeth, no impulse to vigorous action, but in their stead, limp muscles, calm breathing, and a placid face? The present writer, for one, certainly cannot. In like manner of grief: what would it be without its tears, its sobs, its suffocation of the heart, its pang in the breast-bone? A feelingless cognition that certain circumstances are deplorable and nothing more. Every passion in turn tells the same story. A purely disembodied human emotion is a nonentity."

ment" even when these objects appear to be related causally to its emotion. The reaction in other ways also, is not adapted to the surroundings: "the cat will dash itself repeatedly against the sides of the cage and neglect a ready avenue of escape." All the pseudaffective (bodily) reactions "cease abruptly at the end of the stimulus without leaving any of the residue (mewing, trembling, hiding) ordinarily observed after 'true' ('felt') emotional states." The activity induced by hypothalamic stimulation seems to carry no greater emotional connotation than would contraction of muscle induced by stimulation of a motor nerve. Further, if the hypothalamus is stimulated while the animal is behaving normally, its activities continue almost unchanged until mechanically interfered with by the motor components of the induced responses. For instance, during hypothalamic stimulation, an

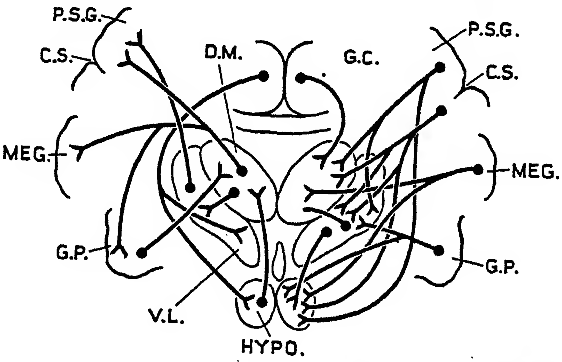


FIG. 428.—Demonstration by the Method of Physiological Neuronography of the Great Wealth of Interconnections between Hypothalamus, Thalamus and Cortex (Cat) which Integrate these Masses into a Functioning Entity. (Murphy and Gellhorn, *J. Neurophysiol.*, 1945, *8*, 440.)

Left: fibres from deep nuclei to cortex.
Right: fibres from cortex to deep nuclei.
GC., cingular gyrus. P.S.G., M.E.G., G.P., various cortical gyri. C.S., cruciate sulcus. D.M., V.L., dorsomedial and ventrolateral nucleus of thalamus. HYPO., hypothalamus.
(There is of course no need to make any attempt to remember the details shown in the Figure.)

animal will often continue as best it can to lap milk, purr, clean its fur, or respond to petting despite the appearance of the typical changes of "sham rage." These results must be contrasted with the immediate abandonment of feeding or purring under the influence of true (complete) rage or fear set up in the cat, for instance, by the sight of a dog.

(2) In man, intravenous injection of adrenaline, which produces intense and generalized sympathetic overactivity, though it gives rise to vague symptoms of disquiet, due in part to the violent impact of the heart beating forcibly and sometimes irregularly against the chest wall, does *not* give rise to an affective state in any way comparable to that experienced in fear, an emotion in which sympathetic overaction is so prominent a feature.

(3) The clinical evidence (p. 663) makes it quite certain that the mental and physiological sides of an emotion can be completely dissociated: a

patient with emotional palsy may feel deeply but displays no bodily changes; a patient with a pyramidal tract lesion may manifest the viscero-somatic changes which ordinarily accompany emotions but in his case the "affect" is lacking (he feels nothing). But though the viscero-somatic changes are not solely or even mainly responsible for the affect, they may modify it to some extent. Afferent impulses from the viscera are normally few (in conditions which do not give rise to pain), but they are presumably set up in a modified form when the state of the viscera is altered; again, afferent impulses from the body wall may be numerous in emotional states. The backlash occurring during the emotional exteriorization may contribute to, or modify, the effective state, but it *reinforces* rather than initiates the affect. (Thus the actor may be really moved by his acting, *i.e.* by voluntary simulated emotional exteriorization.)

More important than the backlash from the body is the complex sustained activity of the higher levels of the central nervous system which is set up by any situation inducing intense emotion. Thus we find that after removal of the cerebral cortex the pseudaffective reactions of anger or defence though violent (as already emphasized) are short lived, and when compared with the responses of the whole animal in a like situation are "simulacra of mere flashes of mimetic passion."

The remaining two possibilities mentioned above, (i) does the emotion felt precede the bodily changes or (ii) do the emotion and the bodily changes occur simultaneously, represent an attempt to draw a distinction where there is little difference. The cortex and subcortex can no longer be considered as relatively independent levels, one higher, the other lower. It is clear that the cortex, especially of the prefrontal lobes, the thalamus, and the hypothalamus, are so closely interconnected that in the intact person they function as an integrated whole. *In emotions, activity develops in large areas of the brain; these set up a discharge to viscera and muscles (the exteriorization), and at the same time in an unknown manner they give rise to the mental state.* The evidence set out on pp. 670 *et seq.* indicates how important is the rôle of the prefrontal cortex and hypothalamus in relation to both these aspects of emotion.

CONNECTIONS AND FUNCTIONS OF THE PREFRONTAL LOBES[1]

The prefrontal lobes are the parts of the cerebral cortex which lie anterior to the excitomotor cortex (areas 4 and 6 and the frontal eyefield region of area 8) and which extend on to the medial aspect of the hemisphere as far back as the anterior end of the corpus callosum and so include the precallosal part of the cingular gyrus (area 24).[2] The prefrontal lobes increase markedly in size as one ascends the phylogenetic scale; but it should be emphasized that the parietal and temporal lobes enlarge even more conspicuously.

Connections of the Prefrontal Lobes.[3]—These are mainly to and fro

[1] Symposium on Frontal Lobes, *Res. Publ. Assoc. nerv. ment. Dis.*, 1948, *27*.

[2] This region is now frequently called the frontal lobes by clinicians; strictly the frontal lobes are the parts of the cerebral cortex anterior to the central sulcus and so include the excitomotor areas.

[3] Le Gros Clark, *Lancet*, 1948, i, 353.

with the thalamus and hypothalamus; connections are also established with other regions of the cerebral cortex.

AFFERENT CONNECTIONS (Fig. 429).—(1) Many fibres pass from the *medial* nucleus of the thalamus to most of the areas of the prefrontal lobes (areas 8, 9, 10, 11, 12 on the lateral and adjacent medial surface; areas 44–47 in the inferior frontal convolution). Groups of cells in this thalamic nucleus project on to circumscribed areas of the cortex. As the medial nucleus of the thalamus receives afferents from the hypothalamus it follows that the impulses that reach the prefrontal lobes via the medial nucleus represent a "resultant" of hypothalamic as well as of thalamic activity.

(2) Fibres from the *anterior* nucleus of the thalamus project on to the precallosal part of the cingular gyrus (area 24). It is known that the hippocampus sends fibres via its efferent tract, the fornix, to end in the mammillary bodies of the hypothalamus whence a new relay transmits the impulses to the anterior thalamic nucleus. Whatever its functions (which are obscure), the hippocampus is ultimately projected on to a prefrontal suppressor area (area 24).

CLOSED-CIRCUIT CONNECTIONS WITH THE THALAMUS.—Like most regions of the cerebral cortex, the prefrontal lobes establish to-and-fro connections with the anterior, medial, and the adjacent so-called "intralaminar" thalamic nuclei; this type of closed circuit, as discussed on p. 617, is responsible for the "resting" electrocorticogram or electroencephalogram.

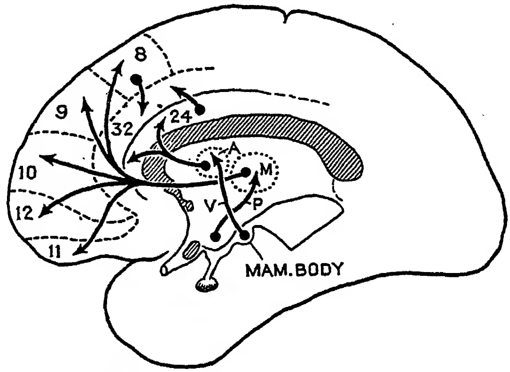


FIG. 429.—Main Afferent Connection of Prefrontal Lobes. (Le Gros Clark, *Lancet*, 1948, i, 354.)

Medial aspect of right cerebral hemisphere.
A = Anterior nucleus of thalamus.
M = Medial (dorsomedial) nucleus of thalamus.
P = Fibres from hypothalamus to M.
V = Mammillothalamic tract.

INTERCORTICAL CONNECTIONS.—(1) Area 32 (on the medial surface, dorsal to the cingular gyrus (area 24)) receives afferents from the frontal suppressor areas 8s and 24s (Fig. 429) and also from the other suppressor areas (4s, 2s, 19s). No other connections of area 32 are known. The significance of this intensive subjection of area 32 to such widespread suppressor influences is not understood.

(2) A long tract runs back from the frontal eyefield in area 8 to area 18 in the occipital lobe (the *parastriate* area, which surrounds the visual cortex and receives afferents from it). The frontal lobes are thus linked with the visual system.

(3) Fibres from the prefrontal areas 44–47 (and also from the occipital area 18) pass into the *temporal lobes*. The prefrontal lobes thus establish connections, both direct and roundabout (*i.e.* via area 18), with the temporal lobes which in their turn receive numerous association fibres from most parts of the cerebral cortex. (Fig. 430.)

Efferent Connections (Fig. 431).—(1) The suppressor bands 8s and 24s discharge to the caudate nucleus (p. 621).

(2) Most of the prefrontal lobe areas project mainly to the *hypothalamus* (and thus indirectly influence the autonomic nervous system and the posterior pituitary) and to the grey matter of the *tegmentum* (midbrain) and *reticular formation of the pons*.

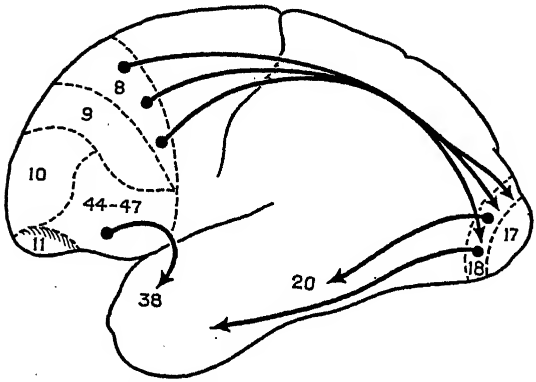


Fig. 430.—Association Tracts connecting Prefrontal Lobes with Occipital and Temporal Lobes. (Le Gros Clark, *Lancet*, 1948, i, 355.)
Lateral aspect of cerebral hemisphere.

As the prefrontal lobes are linked up with the thalamus and the hypothalamus by so many fibres passing in both directions, the whole system may be considered as functioning as an integrated whole. The prefrontal lobes are linked up with the cortical visual areas and the temporal lobes; they also affect the autonomic nervous system via the hypothalamus (and also more directly via the brain stem). The activities of the nervous "complex" consisting of prefrontal lobes, thalamus, and hypothalamus are correlated with some of the *higher intellectual activities*, the *personality*, *emotional affects* and *forms of behaviour of social significance*.

Experimental Studies.—1. Results of Ablation.—Some of the more striking consequences of ablation of the prefrontal areas in monkeys are summarized below.

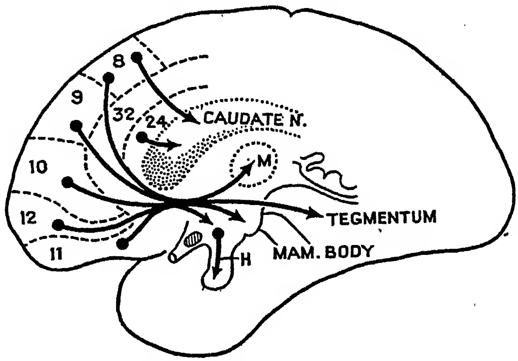


Fig. 431.—Main Efferent Connections of Prefrontal Lobes. (Le Gros Clark, *Lancet*, 1948, i, 355.)
Medial aspect of right cerebral hemisphere.
M = Medial (dorsomedial) nucleus of thalamus.
H = Tract from hypothalamus to posterior pituitary.
Mam. body = Mammillary body.

(1) Alterations of Activity.—In the monkey, prefrontal ablations, especially those involving area 13 (on the orbital surface) (Fig. 432) produce initially a state of *apathy*: the animal sits with a blank expression on its face, the head is drooped, and it stares into space ignoring the approach of human beings; such movements as are carried out are sluggish. After some days or weeks the animal passes into a state of *hyperactivity* which persists unchanged for months. The animal is constantly on the move, incessantly walking

or pacing about "like a caged lion"; the movements are without aim and the animal seems unable to control or check them. Sometimes the movements become almost maniacal in their violence; they cease during the hours of darkness.

(2) Alterations in Emotional Exteriorization.—Monkeys were trained to discriminate between weights; when they chose the heavier weight correctly they were suitably rewarded. The postcentral gyrus was then extirpated, with the result that discrimination was grossly impaired. The animal, finding itself in difficulties about choosing correctly, often flew into a tantrum of rage and annoyance. These tantrums were abolished by lesions experimentally placed in the prefrontal lobes in area 24 (Fig. 431). Cases have been recorded in which transverse lesions of the posterior margins of areas 13 and 14 have produced a clinical state resembling the "sham rage" of the decorticate animal (p. 664).

(3) Alterations in Social Behaviour.—Unilateral or bilateral ablation of the rostral part of the cingular gyrus was immediately followed by marked changes in social behaviour.[1] The monkey lost its pre-operative shyness and fear of man. It would approach the experimenter and examine his fingers with curiosity, instead of cowering (as it would normally) in the far corner of the cage. In a large cage, with other monkeys, it showed no grooming or acts of affection towards its companions; in fact, it behaved as though they were inanimate. It would walk over them, walk on them if they happened to be in the way, and would even sit on them. It would openly take food from its companions and appeared surprised when they retaliated; it was not aggressive, but seemed merely to have lost its "social conscience."

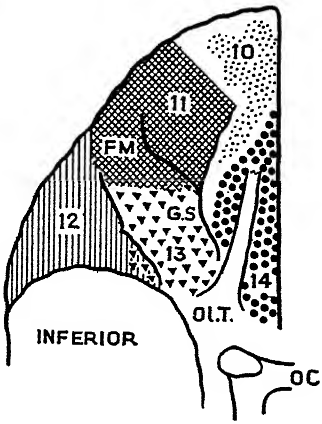


Fig. 432.—Orbital (inferior) Surface of Prefrontal Cortex. (After Ruch and Shenkin, *J. Neurophysiol.*, 1943, *6*, 349.)

Ol.T., olfactory tract
OC, optic chiasma.

(4) Impairment of Memory.—The monkey was shown two inverted cups, and food was placed under one of them; the monkey was trained to raise the cup covering the food; if it chose correctly it was given the food to eat. To test the memory a screen was placed between the monkey and the cups immediately after one had been filled. The normal monkey could still make a correct choice after the cups had been out of sight for 90 seconds; after prefrontal ablation it might choose wrongly after an interval of only 5 seconds.

2. **Results of Stimulation.**—The prefrontal areas by their connections with the hypothalamus and also by their direct brain stem connections can influence autonomic activities. Thus stimulation of area 13 produces changes in heart rate, blood pressure, and breathing (Fig. 433). Stimulation of the central end of the vagus fibres from the lungs produces circumscribed activity

[1] Ward J. *Neurophysiol.*, 1948, *11*, 13. Glees *et al.*, *J. Neurol. Neurosurg. Psychiat.* 1950, *13*, 191.

in area 13 and nowhere else in the cerebral cortex. Stimulation of area 8 affects the motility and secretory activity of the gut.

Effects of Excision of Prefrontal Lobes in Man.—The results in man are very variable and depend in part on the extent and site of the operation and whether it is unilateral or bilateral; graver disturbances are usually noted after bilateral excisions. If the excitomotor areas are spared there is no impairment of volitional movement or alteration in muscle tone or reflexes. Two case reports will be quoted, one in which the effects of the operation were minimal and another in which gross alterations in personality and behaviour appeared.

(1) Case 1 was a mother of six children; she had an extensive unilateral resection of the prefrontal area for tumour.[1] After the operation, "she was careful of her person, and orderly; her sense of humour and insight were intact. She talked intelligently and neither more nor less than good taste demanded." Her powers of expression are illustrated by an extract from a letter written one year after the operation: "We will often think of you at Christmas, and I am thankful I can picture you as you will be, the sunshine in the dining-room and the study, the cheery open fire, the children going out skating, slipping and sliding on the cold crunch of the snow." Her short-comings may be illustrated by the following experience: "Fifteen months after the operation she planned a simple supper for her brother and four members of her own family. She looked forward to it with pleasure and had the whole day for preparation. When the appointed hour arrived she was in the kitchen, the food was all there, one or two things were on the stove, but the salad was not ready, the meat had not been started and she was distressed and confused by her long-continued effort alone. With help, however, the task of preparation was quickly completed and the occasion finally went off quite successfully." It might be added by way of comment that many a woman without a maid might not even have ventured to issue such an invitation, and further, that the experiences described are not unknown when the housewife has both prefrontal lobes intact.

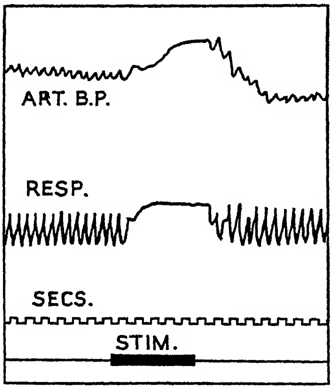


Fig. 433.—Effects of Stimulation of Area 13 on Blood Pressure and Respiration. (After Spencer, *Phil. Trans. roy. Soc. B.*, 1894, *185*, 609.)

Note rise of blood pressure and inhibition of respiration.

(2) Case 2 in whom the larger part of *both* frontal lobes was excised showed impairment of emotional restraint; he was aggressive and was free in his expression of hostility to those around him; though impotent he boasted of his sexual vigour. Thus at the barber's shop: "This is a punk barber's shop." To the barber: "Why don't you go to a barbers' school and learn how to give a good shave." To the next man: "You get the lousiest hair-cut in this place." Or to the doctor: "I will explain to you

[1] Penfield and Evans, *Res. Publ. Assoc. Res. nerv. ment. Dis.*, 1934, *13*, 352.

some other time while I explain the manner in which stock is stocked at the present time. You hear a man offering more than 100 shares of stock, go up to him and say, 'Save me 100 shares'—in other words you are not elected to take this stock until he sells all his stock at a certain price. I think a man should not be elected to that stock until he sells at one-eighth above the price. *Obbligato fortissimo!* Bunk! Do you know what that means?" He is told, "Not in the least," and replies, "Well, it means anything you want it to mean." Dr. Brickner (D) reports the following conversation with his patient (P) in the bathroom. The patient began to scratch his genitalia. D: "What are you doing?" P: "I'm itching myself." D: "It looks to me you are masturbating yourself." P: "Well, I can do that too." To quote the report: "He continued to rub his penis till he had an erection, exhibited same in an off-hand manner to the doctor, and then stopped playing with himself."

Further details of the effects of prefrontal lobe injury in man are given below in the discussion of the results of prefrontal leucotomy.

Prefrontal Leucotomy.—The experimental finding recorded on p. 670 suggested that deliberate lesions of the prefrontal areas might help certain types of mental disorder. The operation of prefrontal leucotomy was introduced by Moniz; a burr hole is made on each side of the skull above the zygoma and behind the orbital margin. A special needle is introduced into the brain and carried through an arc upwards and downwards to cut the subcortical white fibres in the plane of the coronal suture just anterior to the tip of the lateral ventricle. The aim of the operation is to *sever the connections between the thalamus and the prefrontal lobes*. It should be remembered that when the deep thalamocortical connections are cut most of the prefrontal cortex is put out of action as the cortical association fibres are relatively few (p. 669).

The operation just described is a "blind" one, and the surgeon cannot select, precisely, the region which he is to damage or destroy and, in fact, the amount of tissue and the region damaged vary considerably; it is not surprising, therefore, that the reported results differ in patients who showed the same initial clinical disturbances. To an increasing extent therapeutic leucotomies are being carried out by an operation which enables the surgeon to see exactly what he is doing. The newer operations consist of section of selected parts of the corticothalamic connections or of excision of limited areas of the prefrontal cortex (topectomy).[1]

The type of mental disorder for which prefrontal leucotomy has been most helpful is that showing the clinical picture of "mental tension." This is described as a "persistent emotional charge" sustaining and to some extent determining the clinical picture; this "charge" is "always of an unpleasant quality, invariably distressing and sometimes intolerable to the patient. Its presence is shown by irritability, rage, fear or other forms of emotional excitation, insomnia, and on the motor side, restlessness, aggressiveness, destructiveness or impulsive behaviour."

Generally only chronic cases that have not responded to any other form of therapy have been operated on. Following the operation there is usually an initial state of confusion in which the patient has impaired memory, may not know who or where he is and generally shows marked depression of

[1] Beck *et al.*, *J. ment. Sci.*, 1950, *96*, 157.

intellectual activity; there may also be urinary (and occasionally fæcal) incontinence. Some cases pass through a phase of severe irritability in which they are restless, tear at their dressings and become abusive and violent on the slightest provocation; they are usually ravenously hungry. These symptoms pass away gradually; a proportion of the patients (*e.g.* 30% in one carefully selected series) recover sufficiently to leave the mental hospital and live at home or to do useful work; but they rarely fully regain their pre-illness personality and ability.

The principal mental changes after leucotomy are illustrated below.

There is a remarkable change in *mood* to one of "visible happiness" and cheerfulness; one patient reported: "I find the world after all a nice place to live in." After the operation it was said of another patient: "Her outstanding qualities now are her gaiety of spirits, her energy and love of life. The patient is easy going, a good mixer, and fond of social activities. She had a party the other evening and entertained her friends. She played the piano, played duets and games. Maud's friends think that a miracle has happened. She is like the person she was before she became ill, except that she is a much happier edition. A favourite phrase is 'I laughed and laughed' as she relates some joke."

The patient after leucotomy has a high opinion of his own abilities and is talkative and boastful; but in fact he is less conscientious about his work, is less efficient and his *higher intellectual activities are impaired.* He cannot settle down, is easily distracted, and always drifting from job to job. One operated patient described himself as a salesman and supervisor in his father's firm, but the father wrote: "I am sorry he has not found a useful occupation yet; being the son of the director he wanted to be appointed to a managerial post which he cannot possibly undertake. He travels to the city about three days weekly and is helping in the business for about an hour, when he will go off for coffee, return for a little while, and go for lunch, returning usually to his home in the early afternoon." The frequent change of job is partly due to inefficiency and partly to restlessness, though the patient regards it as evidence of ready adaptability. A motor mechanic wrote a year after leaving hospital: "I have got on well since I left. I have been doing nursery work and find it interesting. Before this I worked as a gardener at a hospital here, as a mechanic, also as a sheet-metal worker under a friend of mine, and as a farm labourer." The patients often show a great diversity and multiplicity of hobbies and light interests. "I still have my large flights full of birds. I have forty canaries, also budgerigars and British birds. I keep them mostly for shows. Also I read a lot of books. I have my four hives of bees to look after as well. And I still have my music practice, as I have played a piano accordion for three years now." But there is a marked loss of interest in serious matters. One patient wrote: "A great many strong opinions that I used to have seem to have disappeared altogether. I had, in spite of certain intelligent doubts, great religious faith, but now I seem to have no enthusiasm for any kind of religion." There is something shallow about their character; they lack personal friends. "She has not made any new friends"; "Nobody would be bothered with her"; "Nobody would make a pal of her, she considers herself first always"; "She is not a person you could either like or dislike. She is completely colourless"; "There is a kind of vacancy about her."

The relations may find leucotomized patients *difficult to live with.* The main complaints were (i) shallowness of feeling and lack of affection and of consideration for others, even for children; (ii) domineering manner, self-willed stubbornness, and inaccessibility to reason; and (iii) outbursts of temper and irritability, with no regret for injury done to others. Lack of affection was especially distressing to people who had shown the utmost consideration for the patient returning after a long absence. They found that the patient accepted it as a matter of course. Wives and children felt the emotional estrangement and indifference, and suffered under it. "He is not like his old self; for instance he does not take the interest he should in our little girl, aged three years. He doesn't play with her like other fathers. He cares for no one but himself." Patients do not worry even when they have cause to do so. "He is so far from worrying that he seems to treat without much gravity really serious circumstances. He behaved with his usual calm during a rather unpleasant miscarriage I had in the spring, and attended his badminton club, the cinema, and a fishing excursion just as usual." The patient tends to become a "spiv" type; or as one woman said of her husband: "he is much better but he has lost his soul." In successful cases the symptoms of the previous psychosis (delusions, hallucinations, obsessions, ritual activities such as the endless washing of the hands, sense of guilt and anxiety) more or less disappear.[1]

On the whole there is little irresponsible or perverted behaviour in sexual life, inside or outside marriage, and few reports of serious conflict with the law.

Some patients die weeks or months after the operation from marked wasting or uræmia, resulting it is thought from interference with the cortical control of visceral functions.[2]

CONDITIONED REFLEXES

Pavlov recognized two distinct classes of reflexes.

(i) The *inborn* or *unconditioned* reflex which is present in all normal members of a species: *i.e.* what is generally called a reflex, such as the knee-jerk, light reflex, secretion of saliva when food is introduced into the mouth (Fig. 434), and the flexor reflex.

(ii) The *acquired* or *conditioned* reflex which depends for its appearance on the formation of new functional connections in the central nervous system,

[1] Another striking example of the sort of personality change that may be produced by leucotomy is set out below. Encephalitis lethargica in children, damaging many parts of the brain, especially the basal ganglia, may convert the patient into "a little devil," with uncontrollable and vicious impulses leading to anything from excessively naughty to grossly criminal acts. One patient was taken ill at the age of 10: she subsequently developed temper tantrums, screaming attacks and destructive impulses. She smashed windows and broke crockery, assaulted other children and told lies. Ultimately she became so unmanageable that she was confined to hospital. Bilateral leucotomy was carried out at the age of 31. She subsequently settled down quietly to hospital routine, remained of an even temper, and was cooperative in simple daily tasks; she was contented and could now be allowed home on parole. Her mother reported that she was still somewhat impetuous, self-willed and unable to concentrate for any length of time.

[2] McLardy, *J. Neurol. Neurosurg. Psychiat.*, 1950, *13*, 106.

and is therefore peculiar to the individual. The term "conditioned" refers to the fact that certain conditions must be present if this class of response is to develop. It would help clear thinking if these reactions were called conditioned *responses.*

A simple example will illustrate how a conditioned reflex is established. The introduction of food into the mouth is a stimulus which sets up reflexly the unconditioned response of salivary secretion, and is therefore termed an *unconditioned stimulus.* If a *neutral stimulus, e.g.* the ringing of a bell, is applied so as to coincide with the unconditioned stimulus (the taking of food), and if the procedure is repeated several times the initially neutral stimulus

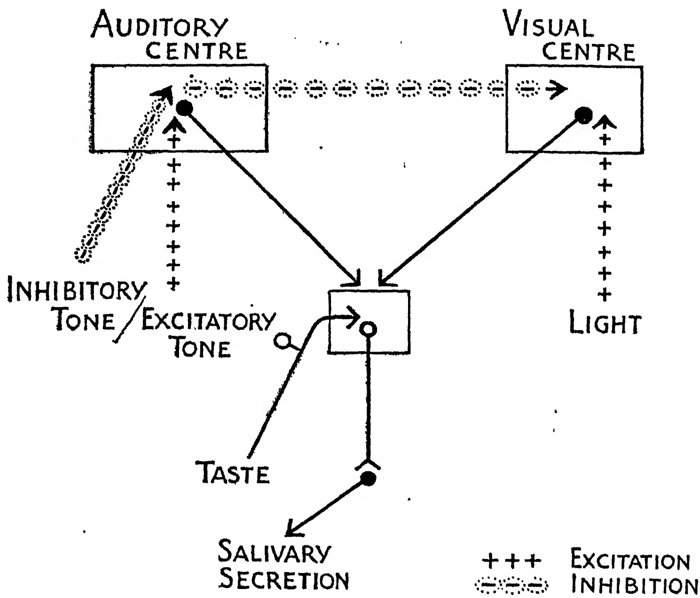


FIG. 434.—Reflex Arcs of Excitatory and Inhibitory Conditioned Reflexes.

Note reflex arc of "unconditioned" salivary secretion in response to taste. The reflex arcs of excitatory conditioned reflexes to a tone and to light are shown. The "inhibitory tone" first inhibits the response to the "excitatory tone" and later by irradiation that to light also.

finally acquires fresh properties (and nervous connections) and can now of itself elicit a secretion of saliva (Fig. 434). The flow of saliva in response to ringing the bell is an example of a conditioned reflex ; the procedure of ringing the bell has become for the individual under experiment a *conditioned stimulus* (*i.e.* one which elicits a conditioned reflex).

Conditioned reflexes are always built up primarily on the basis of inborn reflexes. The salivary reflex is the usual reaction employed because the response of the glands can be expressed quantitatively in terms of the volume of saliva secreted. The salivary duct is usually brought up to the surface, the saliva is collected, and the volume automatically recorded ; to prevent extraneous factors influencing the animal it is usual to have the observer and

the recording apparatus in a separate room from that in which the experimental animal is placed.[1]

Establishment of Positive or Excitatory Conditioned Reflexes.—(1) The animal must be alert and in good health, and there must be complete freedom from all simultaneously operating nervous influences.

(2) The conditioned stimulus (or rather the external stimulus which is to become the conditioned stimulus, *e.g.* for eliciting the salivary flow) must *begin to operate before* the unconditioned stimulus is applied; *e.g.* the bell must *begin* to sound before any food is put into the mouth. [If the reverse order is followed, and the new external stimulus is applied *after* the unconditioned stimulus, it fails to acquire any new properties.] The conditioned stimulus must also be allowed to continue to act so as to overlap the unconditioned, *i.e.* the bell continues to ring while the animal is being fed.

(3) Almost any stimulus if suitably employed may become a conditioned stimulus; it may be one to which the animal was previously indifferent, or even one which is nocuous in character. Nocuous stimuli normally elicit an "unconditioned defence reaction"; but, in spite of this, if they are *not too severe* and are suitably combined with the administration of food they may become conditioned stimuli eliciting a salivary flow, and no movement of defence whatever occurs.

(4) *Necessity for Reinforcement.*—For a conditioned stimulus to retain its new properties it is essential that it should *always* be followed by the unconditioned stimulus. Thus, in the example previously mentioned, if ringing of the bell (conditioned stimulus) is carried out several times alone, and is *not* followed by placing food in the mouth (unconditioned stimulus), it soon ceases to elicit a salivary flow—in other words, the signal has become misleading, it no longer heralds the administration of food and so is ignored (p. 679). The process of following up a conditioned stimulus with the basic unconditioned stimulus is termed *reinforcement*.

(5) The *disappearance* of a natural agency may become an effective conditioned stimulus. Thus a bell is rung continuously while the dog is brought into the experimental room. The sound is then cut out and food is administered. After several repetitions the cessation of the sound becomes an effective conditioned stimulus[2]; and the same applies to a diminution in the loudness of the sound.

(6) *Duration of Time as a Conditioned Stimulus.*—The animal is fed at regular intervals, *e.g.* every 30th minute. After a time it is found that even when the feeding is omitted, salivary secretion occurs at the 30th minute. Or the experiment may be varied as follows: feed the animal every 30th minute, but a few seconds before giving food a metronome is sounded, *i.e.* the conditioned stimulus now consists of the time interval *plus* the sound

[1] The concept of conditioned reflexes is by no means new. Leibniz, in his *Monadology* (1714), writes as follows: "Memory provides the soul with a kind of consecutiveness which resembles reason, but which is to be distinguished from it. Thus we see that when animals have a perception of something which strikes them, and of which they have formerly had a similar perception, they are led by means of representations in their memory to *expect what was combined with the thing in the previous perception*, and they come to have feelings similar to those they had on the previous occasion. For instance, when a stick is shown to dogs, they remember the pain it has caused them and howl and run away."

[2] As happened to Londoners during the Second World War with the "cut-out" of flying bombs.

of the metronome. When the response has become well established, it is found that sounding the metronome at the 29th minute is ineffective, though it gives a complete reaction at the 30th.

(7) *Secondary Conditioned Reflexes.*—These are built up on the basis of an already firmly conditioned reflex, instead of directly on the unconditioned reflex; *e.g.* if a fresh neutral stimulus is suitably applied in conjunction with ringing of the bell (which is already a conditioned stimulus) it finally becomes a conditioned stimulus too.

(8) There is yet a further method of developing conditioned reflexes which is of great clinical interest. Inject apomorphine—this produces in the dog salivation, retching, and vomiting; 1–2 minutes after the injection, sound a tone and continue during the time that the characteristic symptoms make their appearance. After several repetitions, sounding the tone is sufficient to induce the symptoms of the drug, though to a lesser degree. Similarly after several injections of morphine in the dog, the preliminary procedures of sterilizing the needle, cleaning the skin, etc., produce typical symptoms such as nausea, salivation, vomiting, and sleep. In some cases the mere sight of the experimenter is sufficient to produce the same result. It is well known clinically that after a course of injections of hypnotic drugs the injection of saline may have a powerful hypnotic effect; similarly a sleeping draught can be gradually diluted to a considerable degree and yet remain quite effective.

It has often been noted that asthmatics, who develop attacks in the presence, say, of a rose, will be equally affected by an artificial rose. It may be supposed that the appearance (colour, shape) of the rose has become a conditioned stimulus grafted on to the basic stimulus of the chemical constituents of the flower.

The secretion of gastric juice in response to the sight of food is also an example of a conditioned reflex. In puppies brought up on a milk diet, the sight of bread or meat arouses no gastric secretion; it is only after these foods have been consumed on several occasions that the mere sight of them becomes an effective conditioned stimulus for gastric secretion (cf. p. 777).

Conditioned reflexes are mediated by the cerebral cortex; they disappear when the cortex is completely removed. Fig. 434 indicates diagrammatically how the region of the cerebral cortex which receives the impulses set up by the "conditioned" stimulus is believed to establish new functional connections with the basic (inborn) reflex arc.

Conditioned Inhibition.—So far we have been dealing with *positive* conditioned reflexes involving excitation of the nervous system and leading to activity of muscles or glands. We have next to consider phenomena of *inhibition* in connection with conditioned reflexes. These may be divided into two quite distinct groups—*external* and *internal*. In *external inhibition* a positive conditioned reflex is weakened or abolished as a result of some *competing excitatory reaction* which is operating at the same time. By *internal inhibition* is meant that a *stimulus may be made to acquire the power of setting up an inhibitory state in the cerebral cortex.*

1. External Inhibition.—(1) TEMPORARY.—The presence of any *extraneous* stimulus while the conditioned stimulus is being applied prevents the appearance of the conditioned reflex for minutes or even days. The new stimulus arouses the animal's curiosity; the animal looks at and smells

the new object and ignores the conditioned stimulus; but if the disturbing stimulus is repeated on many occasions and thus becomes familiar it ceases to have this inhibitory influence (*i.e.* it is itself ignored).

(2) PERMANENT.—Sometimes the inhibitory influence of the extraneous stimulus is of a more lasting character. Suppose the introduction of weak acid into the mouth has been developed into a conditioned stimulus; if concentrated acid is applied so as to produce severe irritation of the buccal mucous membrane, no conditioned reflex to acid can be obtained until the condition of the mouth is restored to normal ("once bit, twice shy").

2. **Internal Inhibition.**—(1) DIFFERENTIAL INHIBITION.—When a conditioned reflex is successfully established, *e.g.* to a tone of 1000 double vibrations (d.v.) per second, many other tones simultaneously acquire similar properties though to a less degree. This is referred to as the *period of generalization*—when a series of accessory reflexes spontaneously develop around the central reflex. The definitive conditioned stimulus (1000 d.v.) is now regularly reinforced, but the neighbouring tones (*e.g.* 950 d.v.) are never reinforced; the power of *differentiation* rapidly develops, *i.e.* the animal responds strongly to the definitive stimulus but gives no reaction to related stimuli. It can be shown that the unreinforced related stimuli develop *inhibitory* properties as a result of this period of differentiation (Fig. 434):

(i) They show an inhibitory after-effect on the positive conditioned reflex:

12.10. Tone (positive conditioned reflex): 5 drops saliva (reinforced).
12.25. Tone $\frac{1}{8}$ lower (never reinforced): 0 drops saliva.
12.26. Tone: 0·5 drops saliva (reinforced),
i.e. less than reaction at 12.10, showing presence of an inhibitory after-effect.
12.56. Tone: 4 drops saliva (reinforced),
i.e. inhibitory after-effect has nearly passed away.

(ii) Repetition in rapid succession of the unreinforced related stimuli gives rise to a larger inhibitory after-effect (a process of "summation").

(2) EXTINCTION.—It has already been mentioned that if a conditioned stimulus is *not* followed by the unconditioned stimulus—*i.e.* if reinforcement is not carried out—the conditioned reflex is gradually weakened and finally extinguished.

Example: Conditioned stimulus repeated without reinforcement at 2-minute intervals. The salivary reactions in drops were successively: 10, 7, 8, 5, 7, 4, 3.

A conditioned reflex which has thus been extinguished tends to return spontaneously; the *depth* of extinction may be measured by the time which elapses before the reflex returns to its original intensity.

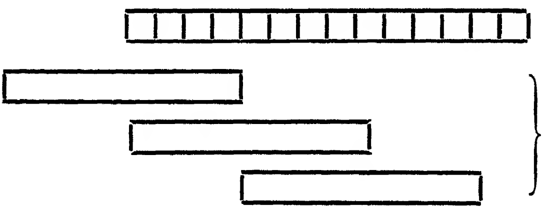
We suggested above that extinction was simply due to the animal ignoring a signal which had repeatedly proved to be misleading. This, however, by no means represents the whole truth; there is evidence that the process of extinction involves in addition the development of an *inhibitory state in the cortical centres.* Thus during the development of extinction, *other* conditioned reflexes which may be present in the animal are also weakened, not only conditioned reflexes built up on the same inborn reflex, but even those based on other inborn reflexes. Further, if the

unreinforced conditioned stimulus is repeatedly applied after extinction is complete, it deepens the state of inhibition in the cortex (cf. p. 683). This is called by Pavlov extinction *beyond the zero*.

(3) Conditioned Inhibition.—The method of inducing conditioned inhibition—*i.e.* a conditioned reflex of an inhibitory or negative character as contrasted with the positive or excitatory conditioned reflexes discussed at first—is as follows:

Establish a positive conditioned reflex in the usual way: let us suppose that the conditioned stimulus is the sound of a metronome. Now apply an *additional stimulus* in conjunction with the conditioned stimulus, carefully observing certain time relationships which will be described immediately; present this combination on many occasions and never reinforce (with the basic unconditioned stimulus), *i.e.* never give food. It is finally found that the combination gives no salivary response, and that the *additional stimulus has acquired the power of inducing an inhibitory state in the cortex*—*i.e.* it is now a *conditioned inhibitor*.

Time Relationship.—To become an inhibitor the additional stimulus must *overlap at some stage* with the conditioned stimulus, though its onset may precede, coincide with, or follow the *onset* of the conditioned stimulus, as shown in the diagram.



It can be demonstrated that the additional stimulus, when presented alone, *can inhibit other positive conditioned stimuli with which it has never previously been combined*, both stimuli based on the same or on other inborn reflexes; it is claimed that it may even *inhibit the inborn reflex itself*.

Experiment: Two *positive* conditioned reflexes have been set up: the conditioned stimuli are (i) the flash of a lamp, (ii) a rotating object—each of which when presented separately induces salivary secretion.

The *rotating object* (only) was then appropriately combined (*supra*) with a tactile stimulus, so that the compound stimulus produced no salivary secretion. The tactile stimulus thus became a conditioned inhibitor. The following observations were then made:

Time.	Stimulus applied for 1 minute.	Saliva in drops.
1.38 p.m.	Rotating object	16
1.50 ,,	Flash of lamp	17
A. 2.14 ,,	*Flash of lamp + tactile stimulus*	2
2.25 ,,	Routine reinforcement of rotating object by feeding	...
B. 2.43 ,,	Rotating object + tactile stimulus	0

Note that either the rotating object or the flash of the lamp presented alone gave a well-marked salivary flow; when the flash was combined *for the first time* with the tactile stimulus (which was the additional stimulus in the inhibitory combination) the secretory response was reduced to only 2 *drops* (line A). Obviously the additional stimulus (the tactile stimulus) has inhibited *another positive conditioned reflex.* In line B we see how the combination of the tactile stimulus with the rotating object gives no flow at all, though the rotating object alone gives a good flow and has just been reinforced.

Such observations prove that the additional stimulus has become a *conditioned inhibitor, i.e.* one that has acquired inhibitory properties.

The inhibitory combination can be shown to exert an *inhibitory after-effect on its own and other conditioned reflexes.*

Experiment: Positive conditioned reflex established: conditioned stimulus (one which produces salivary flow) is rotating object. Rotating object *plus* tone form an inhibitory combination, *i.e.* they give no salivary secretion; the tone is, therefore, the conditioned inhibitor.

The following observations were then made:

Time.	Stimulus during 30 seconds.	Salivary flow (drops).
3.5 p.m.	Rotating object (reinforce)	7
3.26 ,,	Rotating object (reinforce)	6
3.38 ,,	Rotating object + tone	0
3.58 ,,	Rotating object	1
4.10 ,,	Rotating object	2

It will be noted that the inhibitory combination was applied at 3.38, and that at 3.58 and at 4.10 the positive response to the rotating object alone was still very feeble, 1 and 2 drops respectively, instead of 6 or 7 as at first (inhibitory after-effect).[1]

Use of Differentiation in Study of Analysers.—Pavlov used the term *analyser* to refer to the peripheral sense organ and the cortical cells (" centre ") in which the afferent path ends. The analysis of external stimuli is effected partly peripherally (p. 550) and partly centrally. The functional capacity of the various analysers—*e.g.* auditory, visual, tactile, can be investigated by the method of differentiation: a positive conditioned reflex is established and regularly reinforced; closely related stimuli are never reinforced, and one determines whether they develop inhibitory properties—which would be proof of satisfactory differentiation. The following results have been obtained in the *dog*. It by no means follows (in fact, it is most improbable) that exactly similar conditions exist in man.

Acoustic Analyser.—Small variations in *intensity*, and *intervals* of $\frac{1}{8}$ of a tone (*e.g.* difference between 800 and 812 d.v.) can be readily distinguished. The upper range for pitch in the dog (60,000 d.v.) is much higher than in man. The effects of injury to the central or peripheral part of the analyser can be readily studied. Division of the *corpus callosum* results in loss of the power of localizing the side from which a sound comes.

[1] It is, of course, tempting to compare this inhibitory after-effect with the inhibitory after-discharge described for simple reflexes (p. 544).

After removal of both temporal lobes in the dog, auditory conditioned reflexes still persist to some extent and elementary differentiation can still be effected. Though the principal part of the auditory analyser is located in the temporal lobes, extensions of it are dispersed widely over a larger area of the cortex; possibly the outlying parts are only capable of cruder analytical activity (cf. p. 574).

CUTANEOUS ANALYSER.—This analyser in the dog can recognize small differences of *temperature* (1° C.), the precise *site* of stimulation, and the difference between rough and smooth (cf. p. 569).

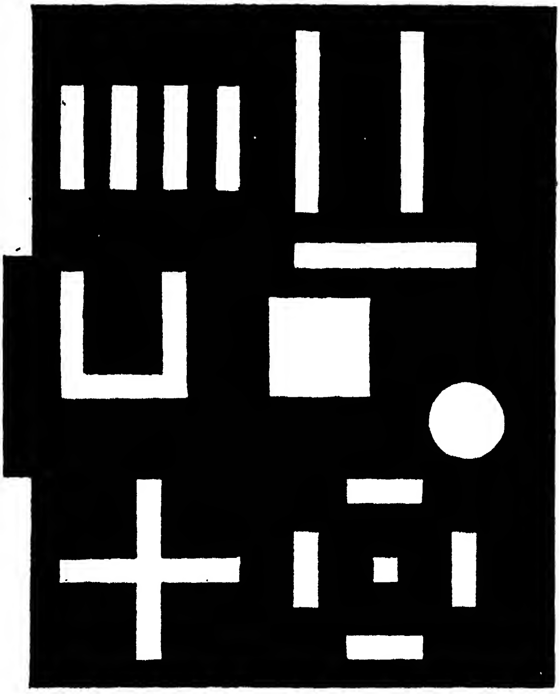


FIG. 435.—Visual Differentiation in the Dog.

The various figures shown were successfully differentiated (by the method of differential inhibition) from the white cross. (Pavlov, *Conditioned Reflexes*, translated by G. V. Anrep. Oxford University Press, 1927.

VISUAL ANALYSER.—Very slight differences in luminosity can be appreciated: different shades of grey can be differentiated which appear identical to the human eye. The colour sense, however, is very rudimentary. A circle can be distinguished from an ellipse, the limit being when the axes of the latter are in the ratio of 9 to 8. The various figures shown in Fig. 435 were successfully differentiated from the white cross.

After extirpation of both occipital lobes *object* vision completely disappears; with limited injuries it is possible to demonstrate loss of parts of the field of vision (cf. p. 580). The interesting observation has been made that

the *anterior* part of the cerebral cortex *in the dog* is capable of rudimentary visual analysis: when the posterior half of the cerebrum is removed, stable conditioned reflexes can still be established to changes in the *intensity of illumination*, and the animal can avoid dark objects in a bright room and walk out through an open door guided by differences in luminosity.

Inhibitory Effect of Prolonged Use of Excitatory Conditioned Reflexes.—When an excitatory conditioned reflex has been elicited on many occasions over a prolonged period the response obtained tends to dwindle and disappear; finally the conditioned stimulus develops *inhibitory* properties and may depress other conditioned reflexes which are present in the animal. If its use is discontinued other conditioned reflexes become enhanced. Sometimes, if an excitatory conditioned stimulus is applied repeatedly during a short period, the animal becomes inert, all positive conditioned reflexes disappear, and the animal may even decline the food which is offered it during routine reinforcement. If new stimuli are now applied, the animal brightens up and the other positive conditioned reflexes return. These observations illustrate well the depressing effect of monotony and the stimulating action of a change in the environment.

Spread of Conditioned Inhibition.—Experiments will be described to show that when conditioned inhibition is produced it first affects small localized areas in one analyser, then it spreads to involve more of the same analyser; later, the effect may involve the greater part or the whole of the cortex, and finally the inhibition may reach the subcortical centres too.

(1) Spread in one Analyser.—Positive conditioned reflexes are established in response to tactile stimulation of four points (1, 2, 3, 4) on the skin of a limb, and a *negative* (*inhibitory*) reflex is established by differential inhibition to a fifth point 0. The spots are thus arranged (distances are measured from 0) (Fig. 436):

0	1	2	3	4
	(3 cm.)	(9 cm.)	(15 cm.)	(22 cm.)

Stimulation of points 1, 2, 3, or 4 yields 5 drops of saliva.

Expt.: (i) Stimulate point 0 three times at 1 minute intervals: 0, 0, 0 drops.
1 minute later stimulate point 1: 0 drops (instead of usual 5).

Pause (ii) Again stimulate point 0 three times as in (i): 0, 0, 0 drops.
1 minute later stimulate point 2: 3 drops (instead of usual 5).

Pause (iii) Again stimulate point 0 three times as in (i): 0, 0, 0 drops.
1 minute later stimulate point 3: 5 drops.

It is clear that the summated inhibitory effect of stimulating point 0 depresses the response from the adjacent spots 1 and 2 but not that from remoter spots, *e.g.* point 3. (See Fig. 436.) One can suppose that the central projection of the spots bears a somewhat similar relationship to that found in the skin (cf. p. 570); this experiment thus demonstrates that inhibition may involve adjacent parts of the analyser without affecting more distant points.

The inhibitory *after-effect* set up by stimulating point 0 disappears in point 1 in 10 minutes, and in point 2 in 5 minutes; *i.e.* recovery occurs first in the outer fringe of the inhibitory effect.

(2) Spread to other Analysers.—When the intensity of the inhibition

is high it may involve other analysers, and conditioned reflexes based on them may be weakened (Fig. 434).

Expt.: Positive conditioned stimulus 1 (tone 4000 d.v.), yields 12 drops. Semitone lower is given inhibitory properties by differentiation. Positive conditioned stimulus 2 (rotating object), based on another—the visual—analyser, yields 7 drops.

(i) Apply tone: 12 drops. Reinforce. Apply semitone twice: 0, 0.
1 minute later, Tone: 5 drops.

(ii) Show rotating object: 7 drops. Reinforce. Apply semitone twice: 0, 0.
1 minute later, Rotating object: 7 drops;

i.e. two repetitions of the semitone produces inhibitory effects on other

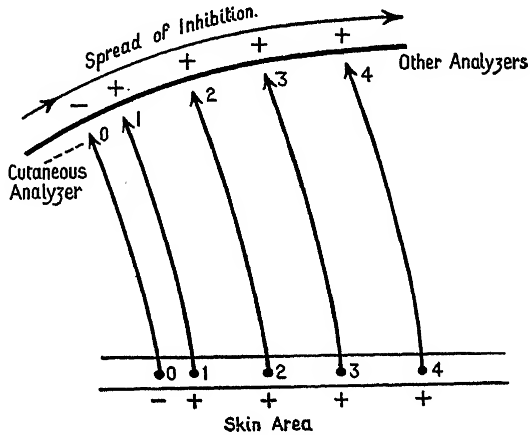


FIG. 436.—Spread of Conditioned Inhibition in the Cerebral Cortex.

0, 1, 2, 3, 4: points on skin and corresponding regions of cutaneous analyser. The sign + (on the skin area) indicates that the stimulus was reinforced, and the sign − that it was not reinforced. In the cortex, + and − represent central excitation and central inhibition respectively.

conditioned reflexes based on the *same* analyser, but *not* on reflexes based on *another* analyser.

(iii) Apply tone: 12 drops. Reinforce. Apply semitone *four* times: 0, 0, 0, 0.
1 minute later, Tone: 1 drop.

(iv) Show rotating object: 7 drops. Reinforce. Apply semitone *four* times: 0, 0, 0, 0.
1 minute later, Rotating object: 3 drops;

i.e. four repetitions of the inhibitory stimulus depress its own (auditory) analyser further, and also exert an inhibitory influence on the visual analyser.

Recovery takes place more rapidly in the distant analyser.

(3) SPREAD TO WHOLE CORTEX.—When inhibitory reactions are being elicited, the following phenomena may be noted: the inhibitory process involves the whole cortex, as proved by disappearance of all conditioned reflexes. Extraneous stimuli, unless exceptionally powerful, fail to elicit any reaction. The animal, nevertheless, "maintains an alert posture: it stands with wide-open immovable eyes, head up, extremities extended, not seeking

support in the loops, remaining motionless sometimes for minutes and sometimes for hours. On changing the position of an extremity it retains the new position. The presentation of food brings about no reaction." Pavlov suggests that the subcortical centres remain at work so that a posture not unlike that of decerebrate rigidity is maintained reflexly, but independently of the cortical centres.

(4) SPREAD TO SUBCORTEX.—Very often when differential inhibition is being carried out the animal becomes drowsy and falls fast asleep. The brain stem centres are now also inhibited as shown by loss of tone, relaxation of skeletal muscles, drooping of the head, sagging limbs, and body hanging limply in the supporting loops.

On the basis of these observations, Pavlov has argued (probably wrongly) that natural *sleep* is merely a *widespread process of internal inhibition* (cf. p. 686). It is interesting to note, however, that sleep may be of a patchy character, leaving one analyser relatively free, as in the sleep of the anxious mother who is readily awakened by the slightest sound from her child.

Cortical inhibition (or "localized sleep") can be restricted to its own analyser by the appropriate application of excitatory conditioned reflexes. Thus a certain tone was turned into a positive conditioned stimulus; various other tones up and down the scale were given inhibitory properties. If the inhibitory stimuli were applied too often, the animal fell into a deep sleep; but if a suitable balance was kept between the excitatory and inhibitory stimuli, the animal remained alert.[1]

Sleep.[2]—It will be convenient to consider at this point the phenomena of sleep and the views which have been expressed as to their causation and significance.

PHENOMENA OF SLEEP.—The outward phenomena of sleep need no description. There is a fall of blood pressure and a decrease in the heart rate; there is no evidence of cerebral anæmia. The respiratory rate may be unchanged or even increased but the total pulmonary ventilation is decreased sufficiently to produce a rise in the CO_2 tension in the alveolar air and arterial blood. The metabolic rate is lowered; thus in a six-year-old boy the hourly heat production was three times higher when he was awake than when he was asleep. Gastric tonus is unaffected, and well-marked hunger contractions may take place. A smaller volume of urine is secreted. Owing to CO_2 retention there is a respiratory acidæmia which is compensated for by the kidney by the excretion of more acid and ammonium salts. Skeletal muscle is markedly relaxed, and in association with this loss of tone the deep reflexes (*e.g.* the knee-jerk) are abolished. Using accurate recording devices it can be shown, however, that the degree of restfulness varies, and periods occur when some degree of muscular activity can be detected. The normal plantar reflex is absent and is replaced by the Babinski response (p. 692): this indicates that the discharge from the motor cortex is in abeyance during sleep. The light reflex can be readily elicited. The axis of the eyes may be deviated in any direction, but the pupils are always small. The position of the body as a

[1] See Pavlov's book, *Conditioned Reflexes: An Investigation of the Physiological Activity of the Cerebral Cortex*, translated and edited by G. V. Anrep, Oxford University Press, 1927, for further particulars; the details of experiments cited in this chapter are all taken from this work. Another point of view is presented in Masserman, *Behaviour and Neurosis*, Chicago, 1943.

[2] Kleitman, *Physiol. Rev.*, 1929, *9*, 624.

whole varies widely during sleep in different individuals and species. Soldiers have been reported to continue standing on parade or marching after falling asleep; horses normally sleep standing on four legs. But as a rule the righting reflexes are lost during sleep. Cortical action potentials cease or are profoundly modified (p. 625).[1]

EFFECTS OF PROLONGED WAKEFULNESS.—The general effects of this condition (familiar to the many who have experienced long periods of aerial bombardment) are common knowledge, but very few precise experiments have been carried out. Subjects have been kept awake for periods varying from 40 to 115 hours. Negligible physical changes were noted, and it is claimed that the mental capabilities were in no way diminished. This last conclusion can hardly be accepted as it is contrary to general experience; probably the tests employed were of too simple and crude a character. Most people who have had inadequate sleep find it difficult to concentrate, are unable to carry out difficult mental activities, and have lost their normal alertness and sense of well-being.

Theories of Sleep.—Literature is full of explanations of the purposes served by sleep: "Sleep that knits up the ravell'd sleave of care, the death of each day's life, sore labour's bath, balm of hurt minds, great nature's second course, chief nourisher in life's feast," and so on. Possibly the poets are right that sleep enables the body to recuperate from the effects of the toils and troubles of the day. But which part of the body needs such a prolonged recovery period? As we have seen, the autonomic nervous system continues to function in much the same way during sleep as during waking hours, though there is possibly lessened sympathetic and greater parasympathetic activity; the vital centres in the medulla oblongata continue in the main to carry out their normal functions. It is the skeletal musculature (with the striking exception of the muscles used in rhythmic respiration) which is reduced to comparative quiescence; sleep essentially involves reduced and modified activity or actual inactivity of the higher parts of the brain and especially of the cerebral cortex. We do not know why the cerebral grey matter should need these long periods of rest, but as pointed out on p. 492, the oxygen consumption of brain tissue is always very high. It may be supposed that the changes which take place in nerve cells are not fully recovered from during waking hours, and that a sort of "debt" is incurred which can only be discharged during periods of sleep. This view is a modification of the once fashionable "hypnotoxin" theory which suggested that a noxious chemical agent accumulates in the brain, and when present in sufficiently high concentration arrests cerebral activity which does not recover again till the alleged toxin is destroyed. Too little is known, however, of the metabolism of brain tissue to make these more than interesting and untested suggestions.

Pavlov's views that sleep represents a state of what may be called "active" cortical inhibition were discussed above (p. 685). Koch (p. 766) showed that afferent impulses from the carotid sinuses may produce a state indistinguishable from sleep. But it is very difficult to see any real resemblance between the methods employed by Pavlov or Koch to produce sleep in experimental

[1] Depth of sleep varies during the sleep period. It reaches a maximum during the first hour, after which it gradually diminishes, to decrease rapidly towards the time of waking.

dogs and the way in which we normally compose ourselves to sleep. Sleep cannot be regarded as an acquired reaction, which conditioned reflexes (by definition) essentially are, in face of the fact that new-born babies sleep for some 23 hours a day and may have to be wakened for their feeds.

It seems possible that the basic condition of the brain is that of sleep from which we have to be actively stirred into wakefulness. Stress must be laid on the fact that to produce an efferent discharge a continual afferent drive is usually needed. The preparations for sleep reduce the sensory stream to a minimum level: the quiet room, the closed eyes, and the relaxation of the muscles all reduce the incoming impulses from the dominant distance receptors, and also from the proprioceptors. We may suppose that in the absence of an adequate afferent stream cerebral activity is suspended.

Relation of Nerve Centres to Sleep.—Developing this theme it has been suggested that sleep may be the result of some "block" at the level, say, of the thalamus, which cuts off the cerebral cortex from afferent impulse streams. There is a good deal of evidence that sleep and wake are related to the state of this level of the brain. Disturbances of the hypothalamic nuclei, for example, may be associated with derangements of sleep (p. 718). In encephalitis lethargica, derangements of sleep are common, either in the form of prolonged periods of sleep or in the inverse form of marked sleeplessness. In these cases Economo found lesions in the grey matter surrounding the third ventricle and the Sylvian aqueduct. He suggests that these regions act in some way as a "sleep centre." It is more likely that disease of this neighbourhood may interrupt essential ascending or descending tracts. Disorders of sleep may occur without apparent involvement of the principal long sensory or motor paths.

Injection of ergotamine (ergotoxin) into the third ventricle of animals may induce sleep of a seemingly normal character. The significance of this observation is difficult to assess. The sleep may be the result of the action of the drug on the adjacent hypothalamic nuclei.

Hess has described a technique by means of which special insulated, fine-pointed electrodes can be introduced into the brain of unanæsthetized normal animals and localized areas may thus be stimulated. If the grey matter in the midline in the region extending backwards from the thalamus along the neuraxis is stimulated, sleep is induced which has all the normal physical accompaniments. When stimulation is discontinued the animal wakes up again. The experiment can be repeated many times in the same subject, proving that it is not due to destructive effects. Stimulation of closely adjacent regions may give rise to bursts of activity—*e.g.* "une veritable gloutonnerie," epileptiform attacks or coordinated defæcation or micturition. Hess suggests that a subcortical centre "presides" over the fundamental cycle of sleep and wake. It might act in one of two ways: by sending afferent inhibitory impulses to the cortex which induce a state of sleep or (as mentioned above) by interrupting excitatory impulses which maintain a state of wakefulness.

No final conclusions can yet be reached, but it seems clear that the same common end result of sleep may be brought about in a variety of ways.

THE SPINAL CORD AND BRAIN STEM[1]

The Spinal Cord (Medulla Spinalis)

We have already discussed fully the various ascending and descending tracts of the spinal cord. It remains here to summarize the facts briefly (Fig. 437).

A. **Descending Tracts.**—(1) Crossed pyramidal tract [lateral corticospinal tract]: lies in the lateral columns of the cord. It is derived from the cells of layer V of the opposite precentral cortex (p. 632).

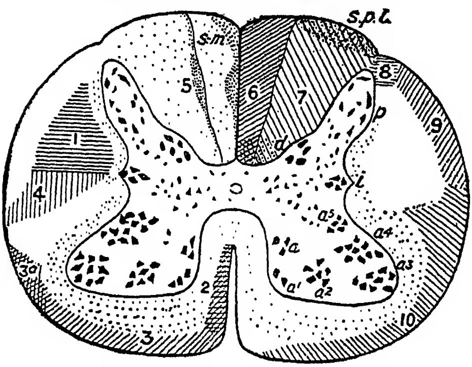


Fig 437.—Diagram showing the Ascending (Right Side) and Descending (Left Side) Tracts of the Spinal Cord. (Sharpey-Schafer.)

1. Crossed pyramidal tract; 2. Direct pyramidal tract; 3. Ventrolateral descending; 3*a*. Bundle of Helweg; 4. Rubrospinal; 5. Comma; 6. Funiculus gracilis (tract of Goll); 7. Funiculus cuneatus (tract of Burdach); 8. Lissauer's tract; 9. Dorsal spinocerebellar tract; 10. Spinothalamic and ventral spinocerebellar; *a.m.* Septo-marginal; *s.p.l.* Superficial dorsolateral fibres; a–a^5 Cells in ventral horn; *i.* Cells in lateral horn of grey matter (connector cells of sympathetic); *p.* Cells of dorsal horn; *d.* Clarke's column of cells. The scattered dots indicate the situation of "endogenous" fibres (arising in the grey matter of the cord) and having for the most part a short course. Many of these which lie near the grey matter have not been indicated.

(2) Direct pyramidal tract [ventral corticospinal tract]: in the ventral columns close to the anterior median fissure. It does not extend below the midthoracic region. The fibres cross in the ventral commissure to the opposite side. The pyramidal tracts end round ventral horn cells both directly and indirectly via the dorsal group of interneurones (p. 633).

(3) Descending fibres from (i) the nuclei of the reticular formation of the pons and medulla—*reticulospinal* fibres; (ii) olivary nucleus—*bulbospinal* tract; (iii) vestibular nuclei—*vestibulospinal* tract; (iv) red nucleus—*rubrospinal* tract. In man most of the fibres from the red nucleus after crossing, connect with the reticular nuclei of the brain stem; the impulses then continue down in the reticulospinal tracts; (v) superior colliculi—*tectospinal* tract which passes down the median longitudinal bundle in the brain stem. All these descending tracts lie intermingled in the spinal cord in the ventrolateral region and end round ventral horn cells both directly and indirectly via the ventral group of interneurones (p. 585).

Some of these descending fibres are a relay of *inhibitory* pathways from the cortex (p. 623), others are a relay of *facilitatory* pathways (p. 629).

B. **Ascending Tracts.**—(1) Funiculus gracilis and funiculus cuneatus (tracts of Goll and Burdach) in the dorsal columns of the cord; their cell bodies are in the dorsal root ganglia of their own side. These tracts end in the nuclei gracilis and cuneatus in the medulla. The funiculus gracilis lies medial to the funiculus cuneatus (p. 559). The tract of Lissauer [dorso-

[1] Ranson and Clark, *Anatomy of Nervous System*, 8th edn., Philadelphia, 1947.

lateral tract] consists of short ascending and descending nociceptive fibres of the dorsal nerve roots (p. 555).

(2) (i) Dorsal spinocerebellar tract arises in Clarke's column of cells (dorsal nucleus) of its own side and passes up the restiform body [inferior peduncle] to end in the cerebellum.

(ii) Ventral spinocerebellar tract arises in Clarke's column of cells of the same and opposite side and enters the cerebellum via the brachium conjunctivum [superior peduncle] (p. 606).

(3) Spinothalamic tract: from dorsal horn cells of the opposite side of the cord. These fibres cross obliquely in the grey commissure to the opposite ventrolateral region of the white matter, and ascend to end in the thalamus (p. 559). In the spinal cord the spinothalamic tract consists of a lateral and a ventral part.

C. **Propriospinal Tracts.**—The fibres arise in cells in the cord and end in the cord. They are very numerous and may be divided into: (i) *short*, connecting adjacent segments; (ii) *long*, connecting more distant segments of the cord, *e.g.* arm and leg "centres" and thus subserving long spinal reflexes, *e.g.* reflexes from the foot affecting the hand or reflexes in the reverse direction. The propriospinal tracts are both ascending and descending; they weld the spinal cord into an integrated whole, in the same way that longer paths integrate the spinal cord with the brain stem.

The Table below gives information of value clinically in the diagnosis of the level of a lesion of the spinal cord.

SEGMENTAL DISTRIBUTION IN THE SPINAL CORD (COLLIER AND ADIE)

Segment.	Muscles.	Reflexes.
C4 . .	Spinati.	
C5 . .	Deltoid, biceps brachii, brachialis, brachioradialis.	Biceps and supinator-jerk.
C6 . .	Pronators of forearm.	Pronator-jerk
C7 . .	Triceps, extensors of wrist and fingers.	Triceps-jerk.
C8 . .	Flexors of wrist and fingers	
Th1 . .	Small muscles of hand.	
Th2–10 .	Intercostal muscles.	
Th7–12 .	Muscles of abdominal wall.	Abdominal reflexes
Th12–L3 .	Psoas.	L2, cremasteric reflex.
L3 . .	Adductors of thigh.	
L4 . .	Quadriceps, abductors of thigh	Knee-jerk.
L5 . .	Hamstrings.	
S1 . .	Glutei, posterior calf muscles.	Ankle-jerk.
S2 . .	Anterior tibials, peronei, small muscles of foot.	Plantar reflex.

Complete Trans-section of Cord.—The account which follows is based on the findings in cases of spinal injury occurring during the First World War;[1] it is supplemented by the somewhat different observations made during the Second World War.[2] The accidents of civil life or acute transverse section following on inflammatory or vascular lesions produce very similar phenomena; practically identical results are seen, too, in the experimental animal.

[1] Head and Riddoch, *Brain*, 1917, *40*, 188–263. Riddoch, *ibid.*, 263–402.
[2] Kuhn, *ibid.*, 1950, *73*, 1.

In *acute* trans-section of the cord the patient feels himself cut in two. The higher centres are unaffected and the mind remains clear, but the whole of the body *below* the level of the lesion is deprived of all activity.

Stage of Flaccidity.—The muscles are completely paralysed; all the reflexes are abolished and muscle tone is lost; the muscles lie in any position imposed on them by gravity. There is complete loss of all sensation; cramp-like pains, however, are present at the level of the lesion. The bladder and rectum are generally paralysed. The sphincter vesicæ, however, frequently retains its functions, or recovers very rapidly, with consequent *retention of urine* (see p. 772). The penis is flaccid, and erection is impossible. As the vasoconstrictor fibres leave the cord between the first thoracic and second lumbar segments, a trans-section below the level of the second lumbar produces very little fall of blood pressure, while section at the first thoracic level causes a fall of blood pressure equal to that resulting from destruction of the vasomotor centre, *i.e.* to a level of 40 mm. Hg (cf. p. 303).

The venous return from the limbs depends largely on the contractions of the skeletal muscles. The amount of blood reaching a muscle depends, too, on the functional activity of the part (p. 432). Owing to the complete immobility of the lower limbs, the blood flow reaching them is decreased and the venous return is impeded. The paralysed parts are consequently cold and blue. The limbs may swell if any abnormal tilting of the pelvis obstructs the femoral veins. The skin is dry and is very liable to be affected by serious sloughing bed-sores.

If the lesion is at the level of the sixth thoracic segment, all impulses coming in from the abdominal viscera are cut off from the brain. Griping sensations or distension of the viscera are not appreciated.

The above phenomena belong to the *stage of flaccidity*. It is obvious that the isolated segments of the spinal cord have lost their power of mediating reflex functions. To this temporary state, the term *spinal shock* is applied.[1] In looking for the cause of this condition the following facts must be noted:

Spinal Shock.—(1) The shock affects the *distal* segments of the cord only and not the segments headward to the injury. The monkey with its cord divided in the thoracic region, goes on looking out of the window and catching flies. The *fall of blood pressure* is not responsible for shock, because the fall is equally marked in the headward part of the animal, which does not suffer at all as a result of the trans-section.

(2) *Operative shock* plays no essential part. The method of trans-section is unimportant; cutting the cord across quickly or tearing it deliberately makes no difference. Furthermore, if the animal is allowed to recover from the shock and a *second trans-section* is then made a few segments lower down, the reflex activities of the distal end of the cord are quite unaffected by the second operation.

(3) The higher the animal is in the scale of development, the more profound and the more lasting is the condition of spinal shock. In the cat it lasts a few moments, in the monkey a few days, while in man it persists for about three weeks. In the higher animals the spinal cord has few truly

[1] It is noteworthy that no such condition of "shock" follows trans-section of the *brain stem* in animals. In some patients with complete spinal trans-section the degree of shock is much less marked than here described; some reflexes may never disappear.

autonomous activities; it is largely dominated by the cortical and subcortical centres which in turn are driven mainly by the great sense organs of the head. As the cord is so dependent on the higher nervous levels, the cord is temporarily "thrown out of gear" when their directing influence is cut off, and some time must elapse before it regains its primitive powers of independent activity. This work supports the view that the simple spinal reflex is a "fiction" and that nearly all so-called spinal reflexes also employ long arcs involving the higher levels of the central nervous system (cf. p. 530).

The general principle which emerges from this discussion is very important. The nervous system does not consist of a series of isolated units but is a closely knit and integrated whole. Damage to any part of the nervous system disturbs its smoothness of working, and until compensation has been established, the functional failure is more severe than can be accounted for by the anatomical lesion. We have already noted that immediately after an acute lesion of the internal capsule in man, producing hemiplegia, reflex spinal activities in the paralysed limbs remain in abeyance for some time, *e.g.* tone is lost, and the reflexes are absent (p. 642). To this depression of function in distant parts of the nervous system the term *diaschisis* (Monakow) is applied.

Stage of Reflex Activity.—As the stage of shock passes off, functional activity returns first in *smooth* muscle. The sphincter vesicæ (if affected at all) recovers very soon, but the detrusor of the bladder regains its powers more slowly. The consequent retention of urine must be dealt with by catheterization. Tone next returns to the hitherto paralysed blood vessels, as the connector cells in the cord begin to act independently of the vasomotor centre. The blood pressure is thus restored to about its normal level. The isolated segments of the cord can also mediate as centres for vasomotor reflexes.

1. **Muscle Tone.**—Tone in skeletal muscle returns after two or three weeks. The flexor muscles of the lower limbs now become less flabby and offer some resistance to the fingers. This returning tone is, of course, reflex in character and is produced by impulses entering the cord from the muscles. It is worth noting that the isolated cord "favours" the *flexor* neurones and muscles, and this fact will make clear many of the findings; thus the extensor muscles remain flabby for a much longer period and never attain the same degree of tone as the flexors. *All* the muscles, however, are *hypotonic*, even the flexors themselves because the stretch reflexes (which are mainly responsible for muscle tone) are feeble when mediated by the spinal cord alone especially in man, and need reinforcement from the brain stem centres (cf. p. 583). The limbs tend to adopt a position of slight flexion, and the paralysis is therefore referred to as *paraplegia in flexion*; the posture is feebly maintained, and the limbs cannot support the body weight. Unless some intercurrent disease is present, the muscles undergo no wasting, because, though paralysed for voluntary movements, they are in constant reflex activity.

2. **Reflex Movements.**—Spontaneous involuntary flexor movements of the limbs occur. The small toes are separated and raised; the anterior tibials, the hamstrings, recti abdominis, and adductors of the thigh can be felt to harden, but their load is initially too great and they fail to move the limbs. Contraction of these flexor groups of muscles is accompanied by reciprocal inhibition of the extensor muscles of the limbs.

(1) FLEXOR REFLEX.—The reflex movement that returns first is the *flexor reflex*. To elicit this reflex a nocuous stimulus is necessary, *i.e.* one which tends to injure the part and which causes pain in the intact organism. The reflex is obtained most easily by stimulating the skin or deep structures of the sole of the foot, but the maximum *receptive field* is much wider than this, so that when the cord has recovered, stimulation of the leg as high as the groin or perineum is effective. From the latter regions a stronger stimulus is necessary, and the response is less constant in its appearance. The movement consists of *dorsiflexion* ("upward movement") *of the big toe*[1] and *abduction of the other toes*; these reflex toe movements constitute the *Babinski response* (p. 645); as the reflex spreads there is dorsiflexion of the foot, flexion of the knee and hip, and abduction of the thigh. The antagonistic muscles are inhibited.

The flexor reflex is a withdrawal or defence reflex which removes the limb from an injurious agency. All the muscles which are contracted in the flexor reflex are called physiological flexors; the antagonists (which contract in the *extensor thrust* or the *crossed extensor reflex* (p. 537)) are physiological extensors. It should particularly be noted that the *dorsi*-flexors of the ankle are the physiological flexors in that region.

The *normal* plantar reflex in man in response to stimulation of the sole of the foot consists of a downward movement, *i.e.* plantar flexion of the toes; it only appears with the development of the pyramidal tracts and replaces the more primitive reflex.[2] In the normal plantar reflex the tensor fasciæ femoris contracts; in the abnormal Babinski response the hamstring muscles generally harden. In spinal animals the flexor reflex often spreads to involve the *extensor* muscles of the opposite limb (the *crossed extensor reflex*). In man this does not usually occur. The reflex is either limited to the same limb or produces *flexion* of the opposite limb.

(2) MASS REFLEX.—In some cases a very widespread reaction is readily elicited by scratching any point on the lower limbs or the anterior abdominal wall below the level of the lesion. The response obtained consists of:

(i) Flexor spasms of both lower extremities and contraction of the anterior abdominal wall.

(ii) Evacuation of the bladder even when its contents may only be half the amount which must normally be present before reflex emptying occurs; this may be partly due to the abdominal compression raising intravesical pressure to threshold level (p. 772).

(iii) Profuse sweating below the level of the lesion. To understand the distribution of the sweating, it must be remembered that the sweat fibres to the head and neck arise from Th1, 2, and those to the arm from Th5–9. With a lesion at the level of Th1, the *whole* body sweats when the mass reflex is obtained, as all the sympathetic fibres leave the cord below the level of the lesion.

[1] The other toes may sometimes plantar flex.

[2] Following *pyramidal tract lesions* the normal plantar reflex is abolished and replaced by a "fractionated" flexor reflex—the Babinski response. In some cases the reflex can be elicited from upper parts of the receptive field, *e.g.* by pinching the calf muscles (Gordon's reflex) or firmly stroking the front of the tibia (Oppenheim's reflex). The Babinski response is the normal finding in new-born infants and may persist during the whole of the first year of life; this may perhaps be related to the supposed non-functioning of the pyramidal tracts at this time.

(3) Coitus Reflex.—This is produced by stimulation of the glans penis, or the skin round the genitals, anterior abdominal wall, or anterior and inner surface of the thighs. The response consists of swelling and stiffening of the penis, withdrawal of the testes because of contraction of the cremaster muscles and curling up of the scrotal skin from the action of the dartos. The recti abdominis, flexors of the hip, and adductors of the thighs also contract. Seminal emission may occur. On ceasing stimulation, the penis and lower limbs relax (cf. p. 1104).

(4) Deep Reflexes.—The *knee-jerk* returns about one to five weeks later than the flexor responses. It consists at first merely of a tightening of the slack quadriceps muscle, but later actual extension of the knee may occur. As pointed out (p. 589), the knee-jerk is a "fractionated stretch reflex"; and stretch reflexes are generally feeble in the spinal animal. It is therefore found that though the quadriceps may contract fairly briskly it relaxes immediately, and the limb drops quite limply and not gradually as in the normal person (cf. p. 610). The ankle-jerk may return later still. If ankle-clonus is present it consists of no more than a few irregular and unequal jerks.

(5) Observations on Cases with Increased Extensor Activity.—The *ultimate* clinical picture in patients studied during the Second World War differed in certain interesting respects from those described above; these differences are probably due to the better general health of the patients, which permitted a maximal degree of functional recovery of the isolated spinal cord. Generally about six months after the occurrence of the trans-section marked activity appeared in the extensor arcs, resulting in heightened extensor reflexes and the appearance of extensor spasms. The detailed findings are summarized below.

(i) The ankle-jerk and knee-jerk became exaggerated; quadriceps clonus and ankle clonus were sometimes noted.

(ii) If the limb muscles were passively stretched abruptly, *e.g.* if the flexed thigh was suddenly extended, a reflex extension of the same or both limbs occurred. The contraction often involved both extensor and flexor muscles, converting the limb into a "solid pillar." Relaxation subsequently developed slowly. A few patients in whom this response was well marked could stand for a long time in a bath of warm water without support.

(iii) The mass reflex as described in (2) above was not obtained. Either mass flexion or mass extension of the limbs occurred, which was *not* accompanied by sweating or emptying of the bladder.

(iv) Stimulation of the glans penis produced the genital response described in (3) above, but it was not accompanied by seminal emission or limb movements.

The spinal trans-section syndrome of the First World War thus seems to represent the functional activity of an *incompletely recovered* spinal cord.

3. Autonomic Reflexes.—(1) Reflex evacuation of the *bladder* is gradually established (cf. p. 772); reflex *defæcation* also occurs.

(2) The skin, which hitherto has been dry and scaly, now shows *sweating* again; it becomes more healthy, and ulcers heal up rapidly. Because of the *improved vascular tone* and the return of reflex activity to the skeletal muscles, the circulation through the limbs is greatly improved, and they become warm and of good colour.

Stage of Failure of Reflex Activity.—If general infection or toxæmia

occurs, failure of reflex function develops. The reflexes become increasingly difficult to elicit; the receptive fields become narrowed down to the optimum areas from which the reflexes can be obtained. The mass reflex disappears. The threshold for all reflexes is raised, and fewer groups of muscles are involved in the motor responses. The muscles waste and become flaccid, and bed-sores develop, which still further lower the general state of the patient. Cystitis commonly develops, and with each recurring attack the bladder diminishes in size and evacuates itself more irregularly and less completely; finally the sphincter vesicæ becomes relaxed and the urine dribbles away.

Incomplete Trans-section of the Spinal Cord.—If the spinal cord is gravely injured, but does not suffer complete division, a state of spinal shock develops identical with that already described. When the stage of reflex activity returns, certain striking differences present themselves. To understand these, it must be remembered that in cases of incomplete trans-section some of the descending fibres, in the ventrolateral columns of the cord (especially the vestibulospinal and reticulospinal tracts) may have escaped injury and so *some connections persist between the brain stem and spinal cord.* As we saw in our studies of decerebrate rigidity (in which the pons and medulla are left in control of the musculature), these levels of the brain stem mainly reinforce the activity of the *extensor* neurones. We, therefore, find in these cases that recovery of functional activity is most marked in the extensor groups of muscles.[1] Though a hard-and-fast distinction will be drawn here between complete and incomplete trans-section of the spinal cord, it can be readily understood that many intermediate and transitional cases occur.

(1) Reflex Tone returns to the extensor muscles, and so the legs lie extended at hip and knee, with the toes pointing slightly downwards. The condition is therefore called *paraplegia in extension.*

(2) Involuntary Movements are relatively infrequent, but when they occur involve an increase of extensor tone, producing downward movements of the feet and toes.

(3) Reflex Movements.—(i) *Extensor Thrust Reflex* (cf. p. 590).—The reflex is elicited as follows: the lower limb is passively flexed and allowed to rest on the bed; the patient's foot is then pressed up with the palm of the hand. Active contraction of the quadriceps and *posterior* calf muscles (the physiological extensors) occurs, and the limb straightens out. This reflex is often absent when the cord is completely divided in man. [In spinal *animals* the extensor thrust reflex is obtained quite regularly, even after complete trans-section, as the spinal cord still retains a considerable degree of control over the extensor arcs.]

(ii) The *flexor reflex* can be obtained by nocuous stimulation of the sole of the foot. The flexion movement, however, is small, and the receptive field only extends to the knee. It is usually accompanied by active and forcible *extension* of the *opposite* limb (crossed extensor reflex).

(iii) Gentle flexion of one limb produces extension of the opposite limb (Phillipson's reflex). The flexed limb then becomes extended and the opposite one flexed; the responses alternate in each limb, producing a *steppage* movement.

We see how the range of reflex response is greater now that more of the

[1] As pointed out on p. 693 in some cases of *complete* spinal trans-section in man, considerable extensor activity occurs.

nervous system is available. It is clear also that movements of locomotion can be carried out to some extent (reflexly) by the lower levels of the central nervous system.

(iv) The extensor *deep reflexes* are slightly easier to elicit than normally. But the most striking feature of the knee-jerk is the prolonged period of relaxation which, as always, accompanies the heightened tonic activity in the quadriceps muscle (cf. Fig. 417, p. 643).

Gradually increasing trans-section of the spinal cord may result from compression. A common sequence of events when pressure on the cord is produced by tumour is (i) *hemisection* of the cord, (ii) *incomplete* trans-section, (iii) *complete* trans-section, occurring in that order. We shall consider the outstanding features of these conditions briefly.

Hemisection of the Spinal Cord (Brown-Séquard Syndrome).—The description of the syndrome is based entirely on the results of clinical observation, and it is customary to assume tacitly that the lesion is of *gradual* development and not of acute onset. The principal phenomena may be classified in the following way:

1. Above the Level of the Lesion.—There are no abnormal signs, except possibly a zone of cutaneous hyperæsthesia in the skin area (on the same side) corresponding to the peripheral distribution of the dorsal nerve root which enters the cord just above the level of the lesion.

2. At the Level of the Lesion.—It is frequently not appreciated that the *lesion* (*clinically*) *usually involves several segments of the cord* which have been compressed or damaged. All nervous elements lying in these segments, or entering or leaving them, may be affected.

(1) Destruction of the ventral horn cells results in a *lower motor neurone type of paralysis* in the corresponding groups of muscles.[1] The features of this type of paralysis must be carefully remembered, as the ventral horn cells may be damaged by many noxious agents, *e.g.* the virus of infantile paralysis, the unknown agent which causes these cells to fall out one by one in progressive muscular atrophy, compression of the cells by the distended cavities of syringomyelia, hæmorrhage within the substance of the cord (hæmatomyelia), or pressure of tumours from outside the cord. Division of the ventral nerve root or of the motor fibres in the peripheral nerves produces similar results.

(i) There is complete loss of all movement in the affected muscles.

(ii) The local reflexes are abolished, including reflex tone, and so the muscles are absolutely flaccid.

(iii) The *structural* changes in the muscle fibres and end plates, the *electrical* changes and the occurrence of *fibrillation* are discussed on p. 505.

(2) Involvement of the sympathetic connector nerve cells may cause vasomotor paralysis in the corresponding segments of the skin.

(3) There is complete anæsthesia for *all* forms of sensation because the dorsal nerve roots are destroyed as they enter the cord. The loss of protective pain sensibility exposes the part to injuries which are unattended to, and so whitlows or ulcers may develop.

3. Below the Level of the Lesion.—(1) *Sensory*.—To understand the distribution of the *sensory* loss it must be remembered that the fibres transmitting impulses for "conscious muscle sense" and some touch fibres ascend

[1] Compare carefully the signs of the clinical upper motor neurone lesion as exemplified by hemiplegia (p. 642).

the dorsal columns of the same side, while the fibres for pain, temperature, and some of those for touch ascend the ventrolateral columns (spinothalamic tract) of the opposite side. Touch sensibility is therefore not lost on either side, but is blunted on both sides. Muscle sense is lost up to the level of the lesion on the same side. The upper limit of the pain and temperature loss on the *opposite* side does *not* correspond to the level of the lesion, because as we saw (p. 559), these fibres cross obliquely in the spinal cord. The level of the lesion will therefore be several segments higher in the cord than would appear from the upper level of pain loss.

(2) *Motor.*—The *motor* phenomena are all on the side of the lesion, and we would expect them to be identical with those found in cases of complete trans-section of the cord. That is not the case, however. The explanation is probably this: the Brown-Séquard syndrome is a *clinical* syndrome, and it is very likely that in these cases all connections with the brain stem on the affected side have not been severed, and some *vestibulospinal and reticulospinal fibres are still unscathed.* The clinical picture is therefore like that seen in cases of *incomplete* trans-section of the cord: muscular paralysis, extension of the limb at knee and ankle, spasticity, exaggerated and tonic deep reflexes (knee- and ankle-jerk), ankle clonus, lost superficial reflexes, Babinski's sign present.

If the compression affects both sides more uniformly, we get paraplegia in extension, as already described (p. 694). The sensory loss first affects dorsal column sensibility (sense of position and passive movement); later temperature, pain, and touch are impaired in that order.

When the trans-section becomes complete, paraplegia in flexion develops (p. 691) as all the descending tracts are now cut. If infection supervenes, all reflex activity in the isolated cord segments is lost, and complete flaccidity is found (cf. p. 693).

The Brain Stem

The brain stem includes the medulla oblongata (spinal bulb), pons, and midbrain, and contains many important structures closely packed together. In our studies of the ascending and descending paths we have already noted many tracts traversing or relaying in the brain stem. We shall here correlate these facts and endeavour to obtain a coherent idea of the anatomy and physiology of these very important regions of the brain. The study of the subject will be greatly facilitated by reference to the diagrams (Figs. 438–444).

A section through the *lower* [1] *third of the medulla* shows the decussation of the pyramidal tracts (Fig. 438). They pass from the ventral aspect of the medulla through the base of the ventral horn of grey matter, and come to lie in the lateral columns where they descend into the spinal cord. The tip of the ventral horn is thus cut off, and is seen as an isolated mass of grey matter (lateral nucleus) which soon disappears. The dorsal columns are wider, and therefore push the dorsal horns of grey matter farther apart. The tip of the dorsal horn projects on the surface of the medulla as the tubercle of Rolando. In its vicinity is the descending root and nucleus of the fifth nerve, which we shall see in most of the sections because it extends

[1] In the description of the sections through the brain stem the term "lower" (below) is used in the sense of *caudal* "upper" (above) is equivalent to *cephalad* or *cranial.*

from the pons to the upper cervical region. The central canal of the spinal cord is still situated centrally, but is now beginning to pass nearer the dorsal surface. On the surface in the ventrolateral region, on each side, are the rubrospinal tract, the dorsal and ventral spinocerebellar tracts (Fig. 383, A), and the spinothalamic tract (Fig. 355).

A section at the level of the *olivary body* reveals these points (Fig. 439). The central canal is now approaching the dorsal surface of the medulla, and is about to open out at the calamus scriptorius into the fourth ventricle. The dorsal columns have been replaced by masses of grey matter—the funiculus gracilis ends in the nucleus gracilis, and the funiculus cuneatus in the nucleus cuneatus. From the small cells constituting these nuclei a new relay of fibres arises. These take several courses: the majority pass forwards to cross the middle line (internal arcuate fibres), and come to lie dorsal to the pyramid as the medial lemniscus. Some fibres pass to the restiform body (inferior cerebellar peduncle) of both sides (Fig. 383, A).

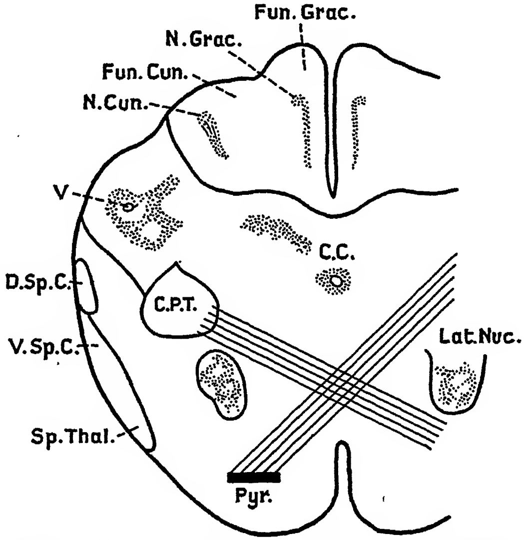


FIG. 438.—Section through Medulla at Level of Decussation of Pyramids (Diagrammatic).

Fun. Grac., Fun. Cun.=Funiculus gracilis and cuneatus; N.Grac., N.Cun.=Nucleus gracilis, Nucleus cuneatus; V.=Descending root of fifth nerve; C.C.=Central canal; Pyr.=Pyramidal tract: C.P.T.=Crossed pyramidal tract; D.Sp.C., V.Sp.C.=Dorsal and ventral spinocerebellar tracts; Sp.Thal.=Spinothalamic tract Lat. nuc.=Lateral nucleus.

In the floor of the fourth ventricle at the level of the calamus scriptorius and in close relation to the dorsal nucleus of the vagus are situated certain important centres, *i.e. cardiac* (p. 270), *vasomotor* (p. 303), *vomiting* (p. 811), and *deglutition* centres (p. 805). The *respiratory* centres lie in the reticular formation of the pons and medulla (p. 384).

The *olive*, which is seen dorsal to the pyramid, consists of a wavy layer of grey matter. Fibres pass to the restiform bodies on both sides. Nothing is known about the functions of the olive; it gives rise to a descending tract, the bulbospinal tract, which lies in the cord in the ventral root zone.

The only change to note in the other tracts is that the fibres of touch in the spinothalamic tract now migrate medially to lie just dorsolaterally to the medial lemniscus, while the fibres carrying pain and temperature remain lateral to the olive (see Fig. 355).

The *nuclei* of the twelfth, eleventh, tenth, ninth, and seventh cranial nerves may conveniently be described here.

XII.—The nucleus of the hypoglossal nerve extends through the lower two-thirds of the medulla. It first lies ventrolateral to the central canal, and when the fourth ventricle appears, it lies in its floor, close to the middle line. The hypoglossal nerve fibres pass out ventrally between the olive and pyramid.

If the hypoglossal nerve is stimulated, the tip of the tongue is pushed over to the opposite side. Conversely, if the twelfth nerve is paralysed, the tongue when projected deviates to the same side (by the unopposed action of the other nerve). The affected side of the tongue shows the ordinary signs of a lower motor neurone paralysis—great wasting of the muscle fibres and consequent wrinkling of the mucous membrane, and fibrillation (cf. p. 505).

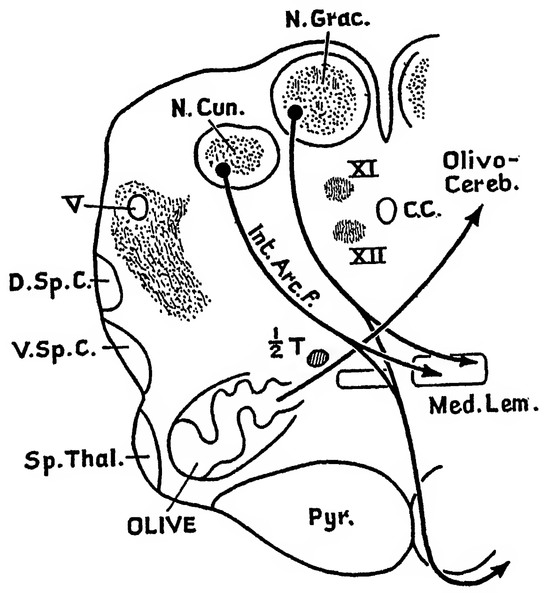


FIG. 439.—Section through Medulla at Level of Olivary Nucleus (Diagrammatic).

Int.Arc.f.=Internal arcuate fibres; Med.Lem.=medial lemniscus; Olivo-Cereb.=Olivocerebellar fibres; ½T.=half touch fibres; XI, XII.=Nuclei of eleventh and twelfth nerves; N.Grac., N.Cun.=Nucleus gracilis and cuneatus; Pyr.=Pyramid; C.C.=Central canal; V.=Descending root and nucleus of fifth nerve; D.Sp.C., V.Sp.C.=Dorsal and ventral spinocerebellar tracts; Sp.Thal.=Spinothalamic tract.

XI.—The nucleus of origin of the spinal part of this nerve consists of cells lying in the lateral part of the ventral horn of grey matter in C1–5. The fibres pass dorsally and then bend outwards to emerge at the side of the cord and medulla as *lateral* roots. These contain large medullated efferent (and some afferent) fibres for the supply of the sternomastoid and trapezius. (The bulbar part of the nucleus (shown in Fig. 439) is best regarded as part of the vagus nucleus.)

X. — The afferent fibres of the vagus arise in the jugular and nodose ganglia from unipolar cells and convey impulses from somatic and visceral structures (p. 712). Some fibres ascend to end in the *dorsal* nucleus, which lies in the floor of the fourth ventricle lateral to the hypoglossal nucleus; other fibres descend to form the tractus *solitarius* and end in adjacent nerve cells (the *nucleus* of the tractus). These two nuclei give rise to the autonomic (involuntary) fibres of the vagus and also connect up with the nucleus ambiguus.

The *nucleus ambiguus* is the somatic ("voluntary") motor nucleus of the vagus, and lies ventrolateral to the dorsal nucleus. The fibres run first dorsally and then curve round to emerge dorsal to the olive and supply such structures as the mucles of the larynx (see Fig. 440).

IX.—The ninth [glossopharyngeal] nerve is arranged in exactly the same

way as the tenth. The afferent fibres have their cell bodies in the petrosal ganglion. They end in nuclei which form the upward continuation of the dorsal nucleus of the tenth (ascending fibres) and of the tractus solitarius (descending fibres). The motor nucleus which supplies the pharynx is in line with the nucleus ambiguus. The autonomic fibres arise in the dorsal nucleus (cf. p. 712).

VII.—Though the seventh [facial] nucleus is situated in the lower pons, it must be considered here because the arrangement is like that already described. The afferent fibres have their cell bodies in the geniculate ganglion and pass in the nervus intermedius to end in nuclei in line with the dorsal nucleus of the ninth and tenth and the fasciculus solitarius. The autonomic fibres arise in the dorsal nucleus (cf. p. 714). The motor nucleus is in line with the nucleus ambiguus. The fibres run dorsally and form a loop round the sixth nerve nucleus and then turn ventrally and laterally.[1]

In lesions of the pyramidal tract (*supranuclear lesion*) voluntary movements in the face are lost, but emotional movements (frowning, smiling) are retained because of the separate motor pathway for emotional exteriorization (p. 642). If the seventh nerve is injured in any part of its course (*infranuclear lesion*) all types of movements of the face are equally affected.

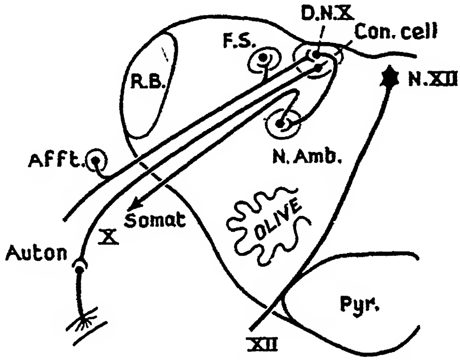


Fig. 440.—Diagram of Mode of Origin of Tenth and Twelfth Nerves.

D.N.X.=Dorsal nucleus X.; F.S.=Fasciculus solitarius Con. Cell = Connector cell which gives rise to autonomic fibres in vagus; N.Amb.=Nucleus ambiguus which gives rise to somatic fibres of vagus.

We may summarize these facts thus:

(i) The motor somatic nucleus of XII. is a column of cells lying near the middle line.

(ii) The afferent fibres of VII., IX., and X. divide into:

(*a*) Ascending fibres which end in a column of cells lying lateral to the nucleus of XII.=column of *dorsal* nuclei of VII., IX., and X.

(*b*) Descending fibres which end in grey matter lying just lateral to the column of dorsal nuclei=*fasciculus* (*tractus*) *tractus solitarius* and its nucleus.

(iii) The autonomic efferent fibres of VII., IX., and X. arise in the column of dorsal nuclei.

(iv) The somatic efferent fibres of VII., IX., X., and XI. arise from a column of cells which extend from the lower border of the pons to C5, and are situated lateral and ventral to the column of dorsal nuclei.

A section at the junction of pons and medulla shows the restiform body and the entrance of the eighth nerve (Fig. 441).

The central connections of the eighth nerve are fully given on pp. 572 and 597. In Fig. 447 the *cochlear* division is seen to pass dorsal to the restiform

[1] It is stated that the fibres supplying the orbicularis oris arise in the twelfth nucleus, and those to the orbicularis palpebrarum in the third nucleus, and are merely conveyed by the facial nerve.

body and the vestibular division ventral. The ventral cochlear nucleus lies between the two divisions of the nerve, and the dorsal nucleus on the dorso-lateral aspect of the restiform body. From the latter the striæ acousticæ cross in the floor of the fourth ventricle, and from the former the trapezoid body traverses the substance of the pons. The two groups of fibres turn up as the lateral lemniscus which is lateral to the medial lemniscus. The *vestibular* division ends in the vestibular nuclei which form an important reflex centre, co-ordinating the position of the eyes and limbs with that of the head and helping to maintain tone in the extensor or antigravity muscles; when these nuclei are destroyed decerebrate rigidity disappears (p. 585).

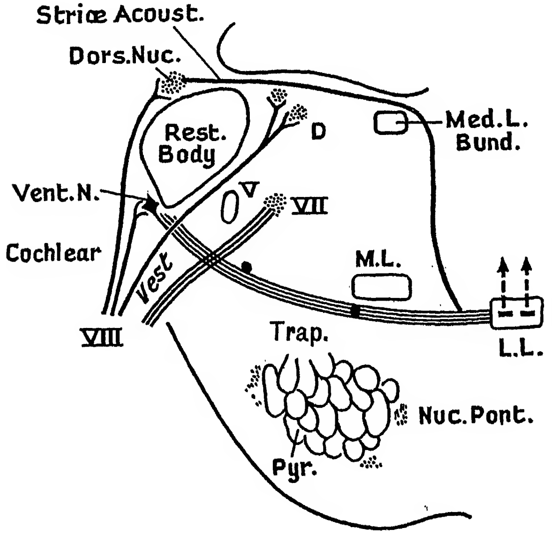


FIG. 441.—Section through the Lower Pons (Diagrammatic).

Striæ Acoust.=Striæ acousticæ; Dors.Nuc., Vent.N.=Dorsal and ventral cochlear nucleus; Cochlear=Cochlear division of VIII; Vest.=Vestibular division; D.=Deiters' [lateral vestibular] nucleus; M.L.=Medial lemniscus. L.L.=Lateral lemniscus; Trap.=Trapezoid body; Med.L.Bund.=Medial longitudinal bundle; Nuc.Pont.=Nuclei pontis; Pyr.=Pyramidal tract; VII=Dorsal nucleus of VII; Rest. body=Restiform body; V=Descending root and nucleus of Vth.

Other features are:

(1) The pyramidal tracts beginning to break up into bundles.

(2) The medial lemniscus, which is now joined by the spino-thalamic tract conveying the fibres for pain, temperature, and touch (Fig. 355).

(3) The restiform body (Fig. 383, A; p. 606).

The Pons.—The appearance of the pons is modified by the presence of numerous transversely crossing bundles of the brachium pontis [middle peduncle of the cerebellum] which are running from the pons to the opposite cerebellar hemisphere and vice versa These break up the pyramidal tract into scattered groups of fibres, between which lie small masses of grey matter, the *nuclei pontis*. The medial lemniscus is now joined by central fibres from the *sensory cranial* nuclei *e.g.* X, IX, VII, V (Fig. 355).

At the upper border of the pons, the fourth ventricle narrows gradually into the aqueductus cerebri [Sylvii]. Above it on each side appear two masses of longitudinally running fibres, the brachia conjunctiva [superior cerebellar peduncles]. These arise *mainly* from the dentate nucleus of the cerebellum, and as they pass forward approach the middle line and decussate to reach the opposite side to end in the red nucleus and thalamus (Fig. 384). The ventral spinocerebellar tract turns over the lateral aspect of this peduncle to enter the vermis of the cerebellum (Fig. 442).

The Midbrain [Mesencephalon].—The midbrain consists essentially of two structures: the superior and inferior colliculi dorsally, and the cerebral peduncles ventrally. Between them is the aqueductus cerebri [Sylvii] which is surrounded ventrally and laterally by the nuclei of origin of the oculomotor nerves (Figs. 443 and 444).

The cerebral peduncles are great masses, chiefly of white matter, uniting the pons with the thalamic region of the cerebrum. They consist of three parts from before backwards:

(1) Basis Pedunculi, containing fibres from the cerebral cortex to the pons and cord. The middle three-fifths are occupied by the pyramidal tracts; the medial one-fifth by the frontopontine and corticonuclear fibres (p. 631), and the lateral one-fifth by the temporopontine fibres.

(2) Substantia Nigra: a mass of deeply pigmented cells, the function of which is uncertain (p. 657).

(3) Tegmentum: a region where longitudinally and transversely running fibres are intermingled in a complex manner. Three decussations take place within it which from below upwards are:

(a) Of the brachia conjunctiva (cf. p. 608).

(b) Of the rubrospinal tracts (cf. p. 688).

(c) Of the tectospinal tracts.

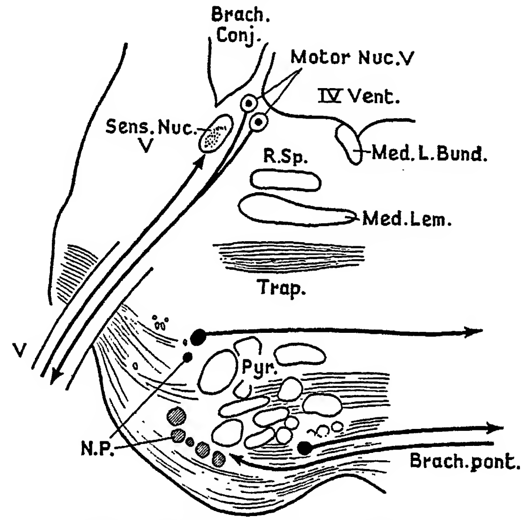


Fig. 442.—Section through Upper Part of Pons (Diagrammatic).

Brach.Conj. = Brachium conjunctivum; V. = Fibres of fifth nerve; IV Vent. = Fourth ventricle; R.Sp. = Rubrospinal tract; Med.Lem. = Medial lemniscus; Trap. = Trapezoid body; Med.L.Bund. = Medial longitudinal bundle; N.P. = Nuclei pontis; Pyr. = Pyramidal tract; Brach. pont. = Brachium pontis.

The Red Nucleus extends from the hypothalamus to the caudal border of the superior colliculi (Fig. 444). It consists of two groups of cells: (i) *n. magnocellularis*—composing the caudal third of the nucleus and made up of large nerve cells which give rise to the rubrospinal tract; (ii) *n. parvocellularis*—which consists of small cells forming the cranial two-thirds of the whole nucleus. In *man*, the large cells are few and the rubrospinal tract is small; most of the cells of the nucleus give rise to a *rubroreticular tract* which ends in the reticular grey matter in the brain stem from which *reticulospinal* fibres transmit the impulses to the spinal cord (p. 593). By means of its afferents from cerebellum, vestibule, and muscles, and its efferent rubrospinal and rubroreticular fibres, the red nucleus plays an important part in *animals* in helping to maintain normal body posture and normal muscle tone. It is the centre for the righting reflexes (Fig. 373) by means of which the body is

restored to its original position after it has been displaced from it. In animals, decerebrate rigidity results from a midcollicular trans-section of the brain stem (p. 583).

The descending pathway from the hypothalamus concerned with emotional exteriorization lies in the tegmentum medial to the red nucleus (p. 666). The lateral lemniscus passes dorsally to end in the inferior colliculi.

The *inferior colliculi* (Fig. 443) were dealt with in discussing the auditory path (p. 572). The *superior colliculi* require further consideration. They are an important centre for visual reflexes. By means of the tectospinal tract they reflexly alter the position of the eyes, head, trunk, and limbs in response to retinal impulses. *Colliculonuclear fibres* pass to the third nerve nucleus to cause constriction of the pupil during the light reflex (p. 578). The tectospinal tract connects the superior colliculi with the pupil-dilator centre in the thoracic cord (Th1 and 2) (p. 708).

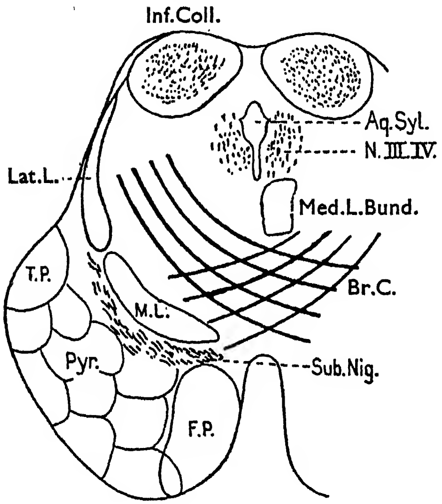


FIG. 443.—Section through Midbrain at Level of Inferior Colliculi (Diagrammatic).

Inf.Coll.=Inferior Colliculi; Lat.L.=Lateral lemniscus; M.L.=Medial lemniscus; Aq.Syl.=Aqueductus cerebri [Sylvii]; Med.L.Bund.=Medial longitudinal bundle; Br.C.=Brachia conjunctiva; N.III.IV.=Nuclei of third and fourth nerves; Sub.Nig.=Substantia nigra; T.P., F.P.=Temporopontine, frontopontine fibres; Pyr.=Pryamidal tract.

RETICULAR FORMATION.—This is found dorsal to the pyramidal tracts in the pons and medulla and in the tegmentum of the midbrain. It consists of small islands of grey matter interspersed with fine bundles of nerve fibres lying in the spaces between the large fibre tracts and nuclei. The functions of the reticular formation have been elucidated to a considerable extent. The *respiratory centres* lie in the reticular formation in the upper pons (*pneumotaxic* centre) and in the upper half of the medulla (*inspiratory* and *expiratory* centres) (p. 384). Reticulospinal fibres connect the brain stem with the spinal motor neurones; these paths are both *facilitatory* and *inhibitory* and serve to transmit impulses from the cerebral cortex, basal ganglia, and hypothalamus to the spinal cord (p. 623, p. 629).

We can now complete our study of the cranial nuclei.

V. [Trigeminal nerve.] — The *efferent fibres* arise from the *motor* nucleus lying at the side of the grey matter bounding the aqueductus cerebri and fourth ventricle in the lower midbrain and upper pons (Fig. 442).

The *afferent fibres* are derived from the semilunar [Gasserian] ganglion and end in a mass of grey matter lying lateral to the motor nucleus—the *principal sensory nucleus.* Long descending fibres are given off which end

round adjacent grey matter.[1] This column of cells and fibres—*descending root* [spinal tract] of the fifth nerve, descends to the upper cervical region of the cord. From these sensory nuclei a new relay of fibres passes across the middle line to join the medial fillet (p. 563).

Oculomotor Nuclei.—Though we speak of the nuclei of the third [oculomotor], fourth [trochlear], and sixth [abducens] nerves as if they were isolated structures, it is important to realize that they simply constitute one long column of grey matter controlling the movements of the eyes. This column extends from the upper part of the midbrain surrounding the ventral and lateral part of the aqueductus cerebri (III., IV.), down to the upper part of the pons in the floor of the fourth ventricle (VI.). The fibres of the third nerve take a curved course in the tegmentum to emerge on the

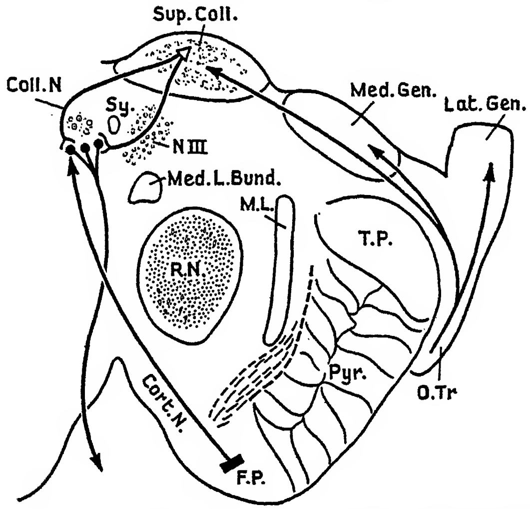


Fig. 444.—Section through the Midbrain at the Level of the Superior Colliculi (Diagrammatic).

Sup.Coll.=Superior Colliculi; Coll.N.=Colliculonuclear fibres; Sy.=Aqueductus cerebri [Sylvii]; Med. Gen., Lat.Gen.=Medial and lateral geniculate bodies; N.III.=Nucleus of third nerve; O.Tr.=Optic tract; Med.L.Bund.=Medial longitudinal bundle; R.N.=Red nucleus; M.L.=Medial lemniscus; T.P.=Temporopontine; Pyr.=Pyramidal tract; F.P.=Frontopontine; Cort.N.=Corticonuclear fibres.

medial side of the crus (Fig. 444). The fourth nerves decussate in the roof of the aqueductus cerebri. The sixth nerve passes between the pyramid bundles to emerge at the lower margin of the pons. Two physiological points must be emphasized:

(1) There are numerous *afferent* fibres in the oculomotor nerves carrying proprioceptive impulses from the external eye muscles. These serve to maintain reflex posture in the muscles, and enable the delicate adjustment of gaze during waking hours to be brought about. The *cell bodies* for these afferent fibres lie on the *trunks* of the oculomotor nerves.

(2) The oculomotor column on *each* side controls synergic eye muscles

[1] Note how the following sensory cranial nerves, V., vestibular division of VIII., IX., and X., end in a principal sensory nucleus and give rise to descending fibres which end in adjacent grey matter.

on *both* sides (cf. p. 637). A lesion of the column results, not in paralysis of an individual muscle, but in loss of a movement in both eyes. Enumerating from before backwards (cranio-caudally), the movements of the eyeballs are controlled by the column thus: (1) pupil-reaction; (2) accommodation (ciliary muscle); (3) upward movements; (4) downward movements; (5) lateral movements of *both* eyes (cf. Fig. 415).

Local Lesions of the Brain Stem.—Owing to the small size of the brain stem, the symptoms are frequently bilateral from a central lesion involving both sides. The points to bear in mind are the origins of the cranial nuclei, the long tracts both motor and sensory, and the cerebellar connections. The long paths to and from the trunk and limbs decussate in the medulla; those connected with the face cross in the upper pons. *Crossed paralysis* and *crossed anæsthesia* are very liable to occur. In the upper pons a lesion may injure the sensory fibres of the fifth, producing anæsthesia of the *same* side of the face, and also involve the medial lemniscus with consequent anæsthesia over the *opposite* half of the body. Similarly, a lesion may destroy the fibres of the facial nerve, giving a lower motor neurone paralysis of the face, and also involve the adjacent pyramidal tract, producing an upper motor neurone paralysis of the opposite side of the body. Involvement of the cerebellar connections produces a varying degree of ataxy and other characteristic signs.

To summarize: the symptoms will be a varying combination of paralysis (of both spastic and flaccid type), anæsthesia, and ataxy. The dilator fibres from the midbrain to the thoracic cord (p. 708) must also be borne in mind. Injury to the brain stem may result in small pupils from cutting off the tonic dilator impulses.

The anatomy and physiology of the *hypothalamus* are discussed on p. 716 where full page references are given.

VI

THE AUTONOMIC NERVOUS SYSTEM

GENERAL ARRANGEMENT AND FUNCTIONS [1]

General Arrangement.—The study of the autonomic nervous system is considerably simplified if attention is first devoted to the broad principles on which the system is constructed. In a somatic spinal reflex arc three neurones may be involved (see Fig. 445): (i) The afferent or *receptor* neurone (A) with its cell *body* in the dorsal root ganglion. (ii) An *internuncial* cell in the dorsal horn of grey matter which by means of its axon transmits the impulse to the ventral horn; it might be called the *connector* neurone (B). (iii) The ventral horn cell and its axon—the *excitor* neurone (C) which transmits the efferent impulses to skeletal muscle (cf. Fig. 325).

In the nerve supply of the viscera three neurones can also be recognized (Fig. 445):

(i) There is an *afferent* neurone proceeding from an internal organ. The nutrient cell lies in the dorsal root ganglion (or its cranial equivalent), and a central process is sent into the grey matter.

[1] Gaskell, *Involuntary Nervous System*, London, 1916. Langley, *Autonomic Nervous System*, Cambridge, 1921. Kuntz, *Autonomic Nervous System* 3rd edn., 1946. White and Smithwick, *Autonomic Nervous System*, 1942. Darrow, *Physiol. Rev.*, 1943, *23*, 1. Gellhorn, *Autonomic Nervous Regulation*, N.Y., 1943.

(ii) The *connector* cell (Gaskell) is situated not in the dorsal horn but in the adjacent grey matter, the exact position differing in the various regions of the nervous system. In the thoracic region, for example, the connector cells lie in the lateral [intermedio-lateral] horn of grey matter. The connector fibre must, of course, connect the afferent neurone with the excitor neurone which actually supplies a viscus. But the excitor cells are not found within the central nervous system; they have migrated outwards to form masses of cells situated peripherally. The connector fibres must, therefore, also wander away from the central nervous system to reach these vagrant groups of excitor cells. The *connector fibres* histologically are *medullated* or *white* fibres; functionally they are B fibres (p. 492); they are called by Langley *preganglionic* fibres, as they pass to a peripheral ganglion.

(iii) The *excitor* cells, as already explained, lie peripherally either as ganglia or as isolated groups of cells. From these cells fibres arise which,

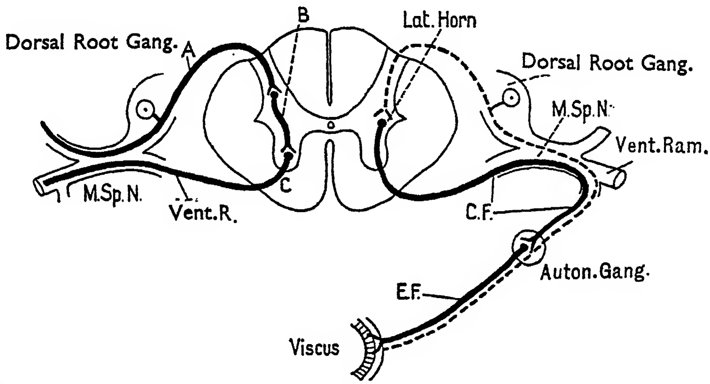


FIG. 445.—General Arrangement of Autonomic Nervous System (on Right) contrasted with that of Somatic Nervous System (on Left). (After Ranson.)

A, B, C=Afferent, connector and excitor neurones of somatic nervous system; Vent.R.=Ventral nerve root; M.Sp.N.=Mixed spinal nerve; Dotted line=Afferent visceral fibre with nutrient cell in dorsal root ganglion; Lat.Horn=Lateral horn of grey matter which gives rise to C.F.=Connector (preganglionic) fibre which passes in ventral root, mixed spinal nerve, and ventral ramus (Vent.Ram.) to end in Auton. Gang.=Autonomic ganglion; E.F.=Excitor (postganglionic) fibre ending in viscus; Dorsal Root Gang.=Dorsal root ganglion.

by devious routes, reach the various organs which they innervate. The *excitor fibres* are grey or *non-medullated*; they are also called *postganglionic* fibres (as they pass from a ganglion to a viscus); functionally they are C fibres (p. 492).

By convention, the term autonomic nervous system is used to include *only the efferent neurones* supplying the viscera, *i.e.* the connector and excitor neurones, using these terms in the sense just defined. The *visceral afferent neurones* are considered on pp. 492 *et seq.*

Connector cells are not present uniformly in all the segments of the nervous system, but in certain regions only, namely (Fig. 446):

(1) In the *brain*, in connection with the nuclei of certain cranial nerves: *i.e.* III., VII., IX., and X.

(2) In the whole of the *thoracic* region, and in the first two *lumbar* segments of the spinal cord.

(3) In the second and third *sacral* segments of the spinal cord.

Connector fibres therefore leave the central nervous system from these regions

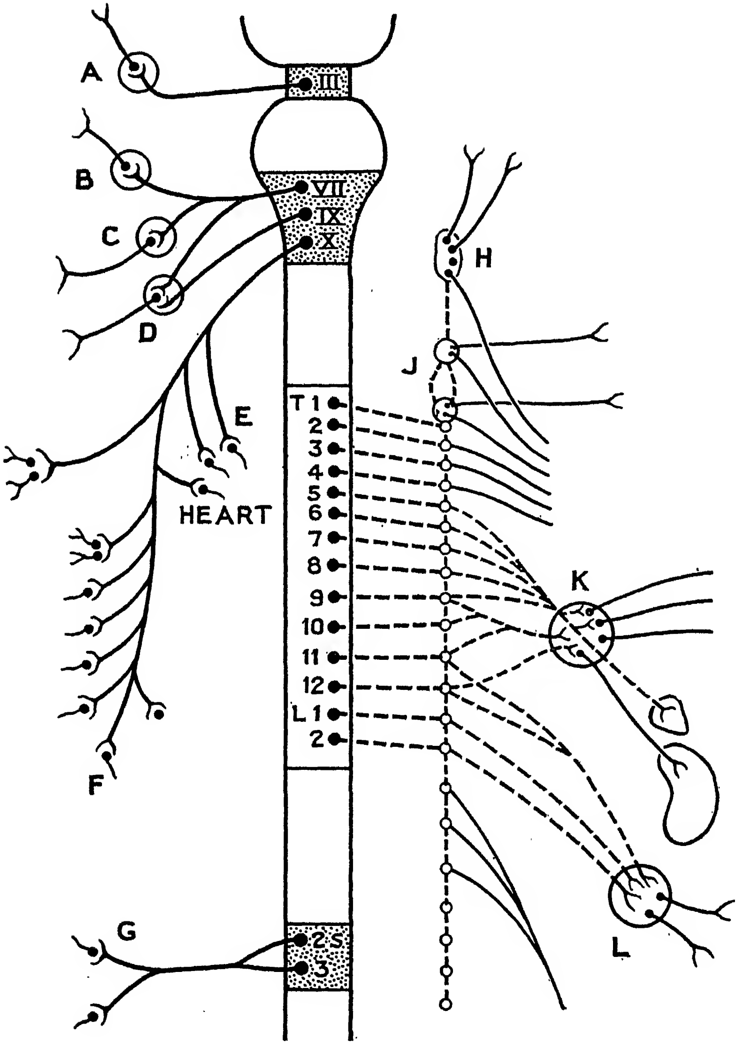


FIG. 446.—General Plan of Autonomic Nervous System.

On *Left*: Cranial and sacral autonomic (parasympathetic) system. Thick lines from III., VII., IX., X., and S2, 3 are preganglionic (connector) fibres. A, ciliary ganglion; B, sphenopalatine ganglion; C, submaxillary and sublingual ganglia; D, otic ganglion; E, vagus excitor cells in nodes of heart; F, vagus excitor cells in wall of bowel; G, sacral autonomic ganglion cells in pelvis; thin lines beyond = postganglionic (excitor) fibres to organs.

On *right*: Sympathetic nervous system. Dotted lines from T1–12, L1, 2 are preganglionic fibres; H, superior cervical ganglion; J, inferior cervical and 1st thoracic ganglia (stellate ganglion); K, cœliac and other abdominal ganglia (note preganglionic fibres directly supplying the adrenal medulla); L, lower abdominal and pelvic sympathetic ganglia; continuous lines beyond = postganglionic fibres.

only, and constitute three great systems of outflowing fibres, which are termed the *cranial* (or bulbar), the *thoracico-lumbar*, and *sacral outflow* respectively.

The autonomic nervous system may be usefully subdivided on the basis of the anatomical situation of the connector cells and fibres. Thus the connector cells in the brain, their axons in the cranial nerves specified, and their related excitor neurones constitute the *cranial division* of the system or the *cranial autonomic*. The connector cells in the thoracico-lumbar region, their axons in the corresponding ventral roots and in their subsequent ramifications and their excitor neurones constitute the *sympathetic nervous system*. The sacral connector cells, their axons in the sacral ventral roots, and all their related excitor neurones constitute the *sacral autonomic* (Fig. 446).

We can therefore make this classification:

AUTONOMIC NERVOUS SYSTEM—

A. Cranial Autonomic.
B. Sympathetic .
C. Sacral Autonomic.

Another classification can be made from a functional standpoint. The cranial and sacral divisions have complementary physiological actions and form the *parasympathetic* system. When stimulated, the sympathetic and parasympathetic nerves produce antagonistic effects on the organs which they *both* supply (*e.g.* heart, pupil). Under *natural* conditions, however, the two systems act *synergistically*, *i.e.* they *cooperate* to achieve the desired end; thus in the case of the heart an increase in rate is due (mainly) to *decreased vagus tone* and (also) to *increased sympathetic tone*.

Some organs are, however, innervated by one division only (*e.g.* uterus, adrenal medulla, and most arterioles from the sympathetic only; glands of stomach and pancreas from the parasympathetic only). The excitor cells of the sympathetic are situated, as a rule, at a distance from the organ innervated; those of the parasympathetic are usually in close proximity to the organs (the sphenopalatine and the otic ganglion (p. 714) are, however, obvious exceptions to the latter statement).

Sympathetic Nervous System.—In describing the anatomical details of the autonomic nervous system it is necessary to note the situation of the connector cells, the path taken by the connector fibres, the position of the excitor cells, and the path of their fibres.

The connector cells of the sympathetic lie in the lateral horn of grey matter in the thoracic region and in the corresponding grey matter in the first and second lumbar segments; the connector fibres pass out in the ventral root corresponding to the segment from which they arise and then enter the mixed spinal nerve. The latter divides into a small dorsal and a large ventral ramus; the connector fibre is continued in the ventral ramus, but soon leaves it to form a branch which passes to the lateral sympathetic chain. This branch is the *white ramus communicans* which passes to the ganglia of the sympathetic; the white ramus is merely a portion of the connector fibre. The connector fibre may end in a ganglion of the lateral sympathetic chain, or pass on to more distantly situated ganglia like those in the neck or abdomen. The excitor (or *postganglionic*) fibres which arise from these ganglia take various routes to the periphery.

1. **Head and Neck.**—The connector cells are situated in the lateral horn of Th1 and 2. The connector fibres pass out in the way just described

to the lateral sympathetic chain, and proceed through the inferior cervical ganglion and the cervical sympathetic trunk, to end in the *superior cervical ganglion*, where the excitor cells are situated (Fig. 446, H). The path of the connector fibres is common to all the structures in the head and neck, but the subsequent course of the fibres varies with the organ supplied.

(1) EYE.—The excitor (grey) fibres to the eye proceed along the coat of the internal carotid artery into the skull and enter the cavernous plexus (in the cavernous sinus). Their further course is not known with certainty, but two alternative (or complementary) routes are described: (i) in the sympathetic root of the ciliary ganglion, through the ganglion itself (without, of course, relaying in it), and along the short ciliary nerves; (ii) on to the semilunar (Gasserian) ganglion, the ophthalmic division of the fifth cranial nerve, its nasociliary branch, and finally in the long ciliary nerves to the eye.

The structures in the eye supplied by the sympathetic are: (i) *dilator pupillæ* muscle; (ii) smooth muscle fibres in the upper and lower lids called respectively the *superior* and *inferior tarsal* muscles which retract the upper and lower lids; (iii) smooth muscle fibres of the *retro-ocular muscle of Müller* which lies in the orbital fascia and which pushes the globe forward; in man Müller's muscle is vestigial and functionless; (iv) *blood vessels.*

Effects on Eye of Stimulation of Cervical Sympathetic.[1]—In all species the pupil is dilated and the lids are retracted producing widening of the palpebral fissure and a "staring" gaze. In man the major effect is on the upper lid, but the lower lid also participates. The effect on the position of the globe is more variable. In anæsthetized dogs, for example, sympathetic stimulation causes the globe to move forwards by 5 mm. (*exophthalmos, proptosis*); but this merely represents a restitution of the *enophthalmos* (sinking in of the globe) produced by the anæsthetic. In anæsthetized man stimulation of the cervical sympathetic at operations produces no *measurable* exophthalmos. It must be remembered that widening of the palpebral fissure gives an illusion of exophthalmos; likewise narrowing of the fissure gives an appearance of enophthalmos (cf. p. 992). The blood vessels are constricted.

Effects on Eye of Section of Cervical Sympathetic (Horner's Syndrome).—The pupil constricts owing to the unopposed action of the fibres of the sphincter pupillæ which are supplied by the third nerve; the palpebral fissure is narrowed mainly owing to drooping (*ptosis*) of the upper lid ("sleepy lid"), but there is also some elevation of the lower lid. There is an illusion of enophthalmos in man, but instrumentally no change in the position of the globe can be detected. The ptosed lid lifts as well as normally during staring showing that its smooth muscle fibres are not used in voluntary movements of the upper lid. The blood vessels are dilated.

There is a centre, situated in the superior colliculi, for the control of the sympathetic fibres to the eye. Stimulation of the appropriate region leads to dilatation of the pupil. The path connecting this midbrain centre with the connector cells in the thoracic cord is the *tectospinal tract* (p. 702). Disturbance of the sympathetic innervation of the eye may thus result from a lesion in the brain stem, the cervical or upper thoracic cord, the upper chest, the neck, the interior of the skull or the orbit.

(2) SKIN STRUCTURES.—For the cutaneous structures of the *head and neck,* excitor fibres pass from the superior cervical ganglion to the four upper

[1] Pochin, *Clin. Sci.*, 1939, *4*, 79.

cervical nerves, and proceed along their cutaneous branches to reach the skin. The skin of the *face* is supplied by excitor fibres which run in the adventitia of the external carotid artery and its branches. The various *glands* of this region also get their sympathetic supply via the blood vessels. This applies to the *buccal*, *parotid*, *sublingual*, *submaxillary*, and *lachrymal* glands, which thus receive vasoconstrictor fibres (and in some species secretory fibres too).

The skin structures (all over the body) supplied by the sympathetic are: (i) *Sweat* glands, with secretory fibres[1] (p. 467). (ii) *Smooth muscle*: arrectores pili, *i.e.* muscle fibres which erect the hairs and produce in man the appearance known as goose-skin; the smooth muscle round the orifices of the body, such as the anus and vagina; the retractor penis muscle. Motor impulses are sent to all of them. (iii) *Blood vessels*: the cutaneous arterioles, veins, and the capillaries receive many vasoconstrictor fibres from the sympathetic (p. 303); some dilator fibres to the skin vessels have also been demonstrated in the sympathetic nerves.

(3) The THYROID receives postganglionic fibres from the middle cervical ganglion; they may control to a *minor* extent thyroid secretory activity (p. 982).

(4) The CEREBRAL VESSELS receive constrictor fibres from the sympathetic along the coats of the internal carotid and vertebral arteries (p. 306).

2. **Thoracic Viscera.**—The connector cells lie in the third and fourth thoracic segments of the cord. The connector fibres in animals pass to the stellate ganglion[2]; in man they end in *all* three ganglia of the cervical sympathetic. The excitor fibres pass from these in the cardiac branches of the sympathetic to reach the *heart*. They increase the force of contraction, the rate, conductivity, and excitability (p. 270). The *bronchi* are dilated by inhibition of the smooth muscle in their walls through fibres having a similar course. The *pulmonary arteries* are *constricted* (p. 307), and the *coronary arteries dilated* by the sympathetic (p. 237).

3. **Fore Limb.**—The connector cells lie chiefly in Th5–9 (and sometimes also in Th2–4) (p. 359). The connector fibres pass to the lateral sympathetic chain to end in the first and second thoracic and inferior and middle cervical ganglia in man (Fig. 446, J). The excitor fibres arise here and join the nerves of the brachial plexus (C5–Th1) constituting their sympathetic roots. The excitor fibres reach the limb via these spinal nerves and are distributed to the *skin* structures already named, to the large arteries (p. 308), to *skeletal muscle vessels*, but probably not to the skeletal muscle fibres themselves (p. 308).

4. **Hind Limb.**—The connector cells lie in Th10–12, L1, 2, chiefly in the lumbar segments. The connector fibres end in the lumbar and sacral ganglia of the sympathetic chain. The excitor fibres arise here and join the lumbo-sacral plexus, and are thus distributed to the lower limb, innervating the same structures as in the arm.

5. **Thoracic and Abdominal Parietes.**—The arrangement in this case is quite simple. The connector fibres pass to the ganglia of the lateral sympathetic chain; they end chiefly in the corresponding ganglia, but also spread up and down the chain to end in adjacent ganglia. Thus one connector

[1] Though the fibres to the sweat glands are anatomically sympathetic they are unctionally *cholinergic*. The responses of the sweat glands to adrenaline and various drugs are considered on p. 481.

[2] This represents the fused first and second thoracic sympathetic ganglia in animals. *In man* the term stellate ganglion is used to describe the *inferior cervical* and *first thoracic* ganglia which are frequently fused together.

fibre may stimulate excitor cells in several ganglia. From each lateral sympathetic ganglion the excitor fibres pass back in the *grey ramus communicans* to the corresponding spinal nerve, and thus to the skin, so that the distribution of the excitor fibres is strictly segmental in character. This is well shown by Langley's studies on the pilomotor fibres. Stimulation of a ventral root (containing the connector fibres) makes the hairs stand up over several segments of skin. On the other hand, stimulation of the grey rami (excitor fibres) affects the hairs of the corresponding spinal segment only.

6. **Abdominal Structures.**—The connector cells for the abdominal organs lie between Th6 and L2. The connector fibres reach the lateral sympathetic chain in the usual way via the white rami. They do *not* relay, however, in the ganglia of the lateral sympathetic chain, but continue through them and leave them as the *splanchnic* nerves, which are therefore still connector or preganglionic fibres. The splanchnic nerves can be divided into two distinct groups: An upper group—the great splanchnic nerves—connects with the ganglia in the upper abdomen, namely, the cœliac (semilunar), superior mesenteric, renal, spermatic, and ovarian ganglia; a lower group, consisting of the lesser splanchnic nerves, ends, in animals, in the inferior mesenteric ganglia (Fig. 446, K, L). In man the arrangement is more complex. The lower connector fibres form the *presacral nerve*, which passes to the *hypogastric* ganglia on the lateral walls of the rectum (Fig. 505). The upper splanchnics arise from Th6–12, and the lower splanchnics from L1, 2 chiefly. From the groups of ganglia mentioned, grey excitor fibres arise which pass mainly along the big arteries to supply the coats of the *blood vessels* themselves and the *organs* which they reach. Thus excitor fibres run with the superior mesenteric artery to the small intestine and part of the large intestine; with the renal artery to the kidney; and along the branches of the cœliac axis to the spleen and liver. The pelvic viscera—*e.g.* the bladder, the uterus, and the distal part of the large intestine—are innervated from the inferior mesenteric ganglia (animals), or hypogastric ganglia in man.

(1) Stomach and Intestines (Fig. 447).—The sympathetic supplies the whole of the small and large intestine and the related sphincters, *i.e.* the ileocolic and the internal anal sphincters. The extent of the sympathetic distribution to the stomach is not so precisely established. It supplies the pyloric sphincter (cf. p. 816), the pyloric region of the stomach, and also the cardiac sphincter (p. 806). It is possible that the sympathetic inhibits movement in the body of the stomach. There are no sympathetic fibres to the œsophagus. Electrical stimulation of the splanchnic nerves results in inhibition of the peristaltic movements of the wall of the gut and diminution of its tone. The sphincter musculature is stimulated and the sphincters become tightly closed.

(2) Blood vessels.—The sympathetic conveys vasomotor fibres to the arterioles constituting the splanchnic area. The vast majority of these fibres are vasoconstrictor in character (though a few vasodilator fibres have been demonstrated). The importance of the vasoconstrictor fibres is very great because the splanchnic area forms the most important part of the peripheral resistance (p. 308).

(3) Bladder.—The detrusor muscle is inhibited, and the sphincter vesicæ and trigonal region are contracted by the sympathetic, the arrangement being analogous to that found in the case of the gut. (For full details see p. 767.)

Sympathetic stimulation in man causes contraction of the muscle coat of the epididymis, ejaculatory ducts, seminal vesicles, and prostate, with resulting ejaculation of semen (p. 767).

(4) The *ureters*, *uterus*, *Fallopian tubes*, and *vas deferens* receive both motor and inhibitory fibres from the sympathetic. None of these structures, be it noted, is supplied by the sacral autonomic.

(5) Secretory fibres are supplied to the *adrenal medulla* to regulate the discharge of adrenaline. (The nervous control of the adrenal gland is discussed on pp. 730 *et seq.*) The cortex of this gland has no nerve supply.

(6) Impulses from the sympathetic stimulate the conversion of the glycogen stores of the *liver* into glucose (p. 857), which passes into the systemic circulation.

(7) Motor fibres are sent to the *spleen*, causing it to contract and discharge the red corpuscles stored in it (cf. p. 226), and also to the *gall-bladder* (p. 801).

(8) Vasoconstrictor fibres supply the *renal* vessels (p. 27).

Organ	Site of Connector Cells.	Site of Excitor Cells.	Route of Excitor (Post-Ganglionic) Fibres.
Head and neck .	Th1, 2	Superior cervical ganglion	
(1) Eye . .	,,	,,	Along internal carotid artery.
(2) Face . .	,,	,,	Along external carotid artery.
(3) Skin of head and neck	,,	,,	With cervical plexus.
(4) Cerebral vessels	,,	Superior and inferior cervical ganglia	Along internal carotid and vertebral arteries.
Thoracic viscera .	Th3, 4	Superior, middle, and inferior cervical ganglia (man) Stellate ganglion (animals)	Cardiac branches of sympathetic.
Fore limb . .	Th5–9 (sometimes also Th2–4)	Middle, inferior cervical, first and second thoracic ganglia (man) Stellate ganglion (animals)	With brachial plexus.
Hind limb . .	Th10–L2	Lumbar and sacral ganglia	With lumbo-sacral plexus.
Abdomen . .	Th6–L2		
(i) Viscera of abdomen proper	Th6–12 (chiefly)	Upper abdominal ganglia (superior mesenteric, cœliac, etc.)	Along blood vessels.
(ii) Pelvic viscera .	L1, 2 (chiefly)	Inferior mesenteric ganglia (animals) Hypogastric ganglia (man)	Along blood vessels and in hypogastric nerves.
Thoracic and abdominal parietes	Th1–12	Ganglia of lateral sympathetic chain	With intercostal nerves.

The table on p. 711 brings out the salient anatomical features of the sympathetic nervous system.

Parasympathetic Nervous System.—This consists of sacral and cranial divisions.

Sacral Autonomic System.[1]—The connector cells lie in the second and third segments of the sacral cord, in the lateral region of the grey matter. The connector fibres pass out in the corresponding ventral roots and then leave these roots to unite to form a single nerve on each side called the pelvic nerves or *nervi erigentes*. They proceed into the pelvis to reach the excitor cells, which are found in the vicinity, or *in the substance*, of the organs they will innervate. In man they relay in the hypogastric ganglia on the lateral walls of the rectum. The structures supplied are the *bladder*, *prostate*, most of the *large intestine*, and the blood vessels of the *penis*. It is not certain how much of the large bowel is thus innervated; it is usually said that the sacral autonomic supplies the whole of the colon and rectum, with the possible exception of the cæcum or proximal colon.

FUNCTIONS.—The sacral autonomic supplies motor fibres to the bladder (p. 767), the large intestine, and to certain muscle fibres round the prostate. It stimulates the detrusor muscle of the bladder (thus emptying that organ), and, as might be expected, it inhibits the antagonist of the detrusor, namely, the sphincter vesicæ internus. Similarly, the colon and rectum are made to contract and the sphincter ani internus is inhibited.

As the name nervi erigentes indicates, erection of the penis is produced (p. 1104).

Cranial Autonomic System.[2]—It is best to describe first the arrangement of the autonomic fibres connected with the vagus.

Vagus.—The *afferent* neurones of the tenth nerve have their cell bodies in the nodose and jugular ganglia. The central axons enter the medulla between the restiform body and the olive to end in the *dorsal* nucleus of the vagus (and the nucleus of the tractus solitarius) (Fig. 440). The dorsal nucleus is situated in the floor of the fourth ventricle, lateral to the nucleus of the twelfth, and contains connector cells for both somatic and autonomic systems. In other words, it contains the homologues of both "dorsal horn" and "lateral horn" cells. Certain large cells give rise to fibres ending in the nucleus ambiguus (or somatic motor nucleus of the tenth), which supplies certain voluntary structures, like the muscles of the larynx. In the dorsal nucleus are found, too, characteristically small connector cells, referred to as the *nucleus intercalatus* (Staderini). These cells give rise to connector fibres which pass *straight out* in the vagus trunk and wander with this nerve (still as connector fibres) to reach various viscera, where they end in excitor cells lying *in the substance* of the organs. It is clear, then that the connector fibres of the vagus correspond to the white rami communicantes or preganglionic fibres of the sympathetic. The grey excitor fibres pass direct from their parent cells to the organ innervated (*e.g.* Fig. 447).

(1) HEART.—The excitor cells lie in the sino-auricular and auriculoventricular nodes. The excitor fibres supply chiefly the junctional tissue of the heart (p. 234), *i.e.* the nodes, the bundle of His and its ramifications—and to a lesser extent the musculature of auricle and ventricle. The vagus

[1] Sacral division of the parasympathetic.
[2] Cranial division of the parasympathetic.

depresses all the activities of the heart, *i.e.* impulse formation, excitability, conductivity, and contractility, with the result that the heart is slowed and its force of contraction diminished (p. 268). The vagus also supplies *constrictor* fibres to the *coronary* arteries (p. 237).

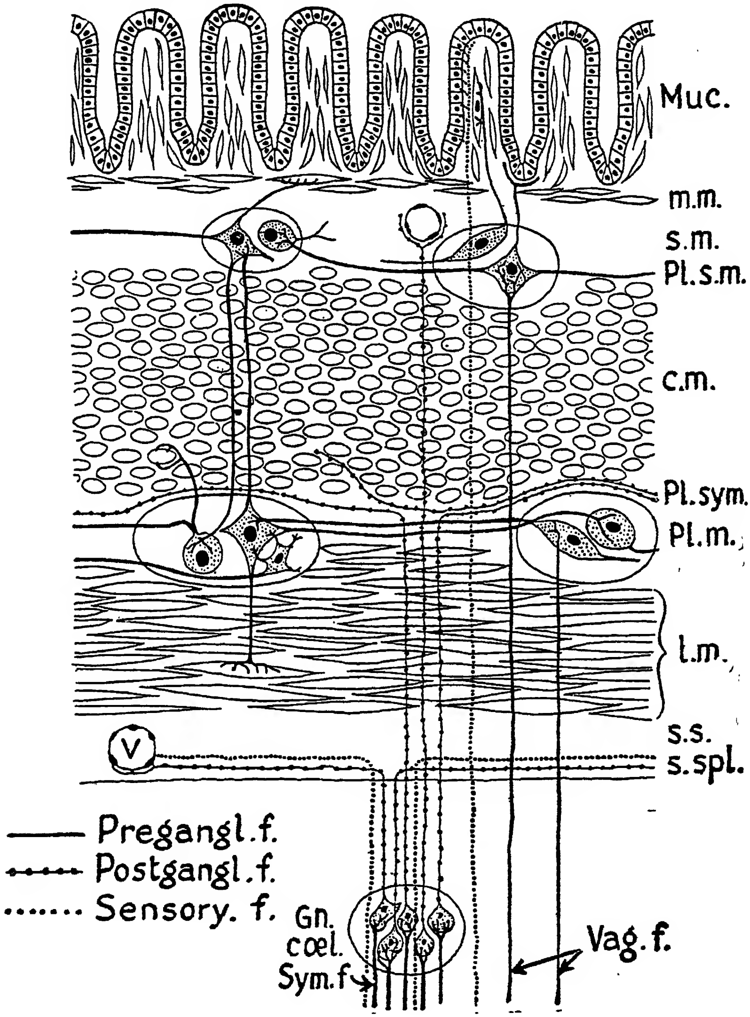


FIG. 447.—Autonomic Innervation of Alimentary Canal. (After C. J. Hill, from Maximow and Bloom, *Text Book of Histology*, W. B. Saunders & Co.

Longitudinal section of gut wall. *Muc.*, mucosa; *m.m.*, muscularis mucosa; *s.m.*, submucosa; *Pl.s.m*, Meissner's plexus [submucous plexus]; *c.m.*, circular muscle; *Pl.sym.*, plexus of postganglionic sympathetic fibres; *Pl.m.*, Auerbach's plexus [myenteric plexus]; *l.m.*, longitudinal muscle; *s.s.* subserosa; *s.spl.*, sympathetic postganglionic nerve fibres; *Gn.cœl.*, cœliac ganglion; *Sym.f.*, preganglionic splanchnic nerve fibres; *Vag.f.*, preganglionic vagus fibres; (cf. Fig. 135) *V.*, blood vessel.

(2) Constrictor fibres are sent to the smooth muscle in the walls of the BRONCHI (p. 409).

(3) ALIMENTARY CANAL (Fig. 447).—The vagus supplies the whole of the alimentary canal from the œsophagus down to the cæcum. Some authorities (Gaskell) believe that the vagus ends at the ileo-colic sphincter, others that it may extend farther. In general, the vagus is motor to the muscle of the alimentary canal (pp. 809, 812, 814) and relaxes the sphincters. Auerbach's [myenteric] plexus (which lies between the circular and longitudinal muscle coats) acts as the excitor neurones of the vagus to the muscle coat of the intestine ; in the case of the œsophagus the corresponding plexus lies on the *surface* (p. 806). The vagus fibres also relay in Meissner's [submucous] plexus in the submucous coat whence excitor fibres pass to the mucous membrane (Fig. 447 ; cf. Fig. 135, p. 234). Important *secretory* fibres pass in the vagus to the glands of the *stomach* (p. 777) and the externally secreting alveoli of the *pancreas* (p. 790). The *islets* of Langerhans also receive a secretory supply from the vagus (p. 917).

Vagal fibres have been traced to the gall-bladder (p. 801), liver (p. 800), and kidney, the excitor cells lying in the organs named ; their significance is unknown.

NINTH NERVE.—If we trace the nucleus ambiguus upwards, we find the *motor ventral nucleus* of the ninth which supplies the stylopharyngeus and superior and middle constrictors of the pharynx. The *dorsal nucleus* of the ninth is continuous with the corresponding nucleus of the tenth and, like it, is a double structure—a fusion of " dorsal horn " and " lateral horn " cells. The small connector cells of the dorsal nucleus give rise to fibres which pass by way of the tympanic nerve and the small superficial petrosal nerve to the *otic ganglion* which constitutes the excitor ganglion. The grey excitor fibres arise here, and join the auriculotemporal nerve to reach the *parotid* gland and supply it with secretory and vasodilator fibres. The dorsal nucleus of the ninth for this reason is sometimes called the *inferior salivary nucleus*. It is worth recalling (p. 581) that taste fibres from the posterior third of the tongue enter the brain stem along the glossopharyngeal to end in this nucleus. Thus afferent impulses from the mouth can readily produce a reflex flow of saliva (p. 774).

SEVENTH NERVE.—The somatic motor nucleus of the seventh is in line with the corresponding nuclei of the ninth and tenth. There is a *dorsal nucleus* containing connector cells (continuous with the column of dorsal nuclei already described), and referred to as the *superior salivary nucleus*. The connector fibres issue in the nervus intermedius, join the facial nerve, and go by way of the great superficial petrosal nerve to the *sphenopalatine ganglion*. The excitor fibres arise here to pass to the *lachrymal* gland by way of the zygomatic nerve, and to the unstriped muscle, vessels, and gland in the *palate* and *nasopharynx* via the palatine nerves. Other connector fibres leave the facial nerve in the chorda tympani to reach the lingual nerve and end round ganglia in the vicinity of the *submaxillary* and *sublingual* glands. From these cells secretory and vasodilator fibres to the glands arise.

As the nervus intermedius conveys taste fibres from the anterior two-thirds of the tongue to the superior salivary nucleus (p. 581), another simple reflex arc for the secretion of saliva becomes evident.

THIRD NERVE.—The connector cells lie in the most cranial part of the oculomotor nucleus in the floor of the aqueductus cerebri. The connector fibres which pass out in the third nerve are often referred to as the *midbrain* outflow in contrast with the bulbar outflow in VII., IX., and X. They leave the third nerve to end in the *ciliary ganglion*, from which excitor fibres pass in the short ciliary nerves to the *ciliary muscle* and *sphincter pupillæ*. The connector cells for the sphincter of the pupil lie in the midbrain, cranial to those for the ciliary muscle. As stated on p. 708, the short ciliary nerves also contain sympathetic grey rami to the dilator pupillæ.

Nicotine Method for Determining Site of Relay Stations in Autonomic Nervous System.—Nicotine has been much used to investigate the path of autonomic fibres. In large doses it *paralyses* the excitor cells of the *whole* autonomic system. After painting a *sympathetic or parasympathetic* ganglion with nicotine, stimulation of the white rami (connector fibres) which end in it is without effect, while stimulation of the grey rami (excitor fibres) produce s the customary results. Stimulation of connector fibres which *pass through* the nicotinized ganglion (without relaying in it) gives rise to the normal responses.

Extirpation of the Sympathetic System.—If the entire sympathetic nervous system is removed in the cat or dog, good health is maintained; reproduction and lactation occur normally in the female. The blood pressure shows an initial fall but most surprisingly full recovery takes place later. It seems that in these animals, even in the absence of sympathetic influences and of secreted adrenaline, the walls of the arterioles develop a sufficient degree of tone to maintain the peripheral resistance. [In *man*, on the other hand, localized sympathectomy leads to vasodilatation which persists for many years (p. 359).] There are no persistent signs of the expected overaction of the parasympathetic; *e.g.* the heart rate and size of pupil soon settle down to about normal. The metabolic rate is lowered by not more than 10 per cent. With emotional excitement there is no change in the blood sugar, no increase in the red cell count, or marked rise of blood pressure such as occur normally (cf. p. 661). The animals are very sensitive to cold, and lose heat more rapidly than normals in a frigid environment. The animals are able to lead a placid existence quite efficiently, but are stated

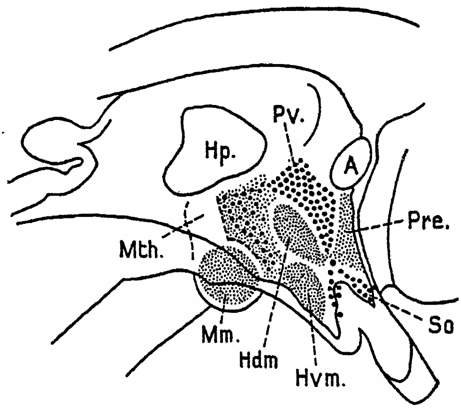


FIG. 448.—Nuclei of Human Hypothalamus. (Le Gros Clark, *J. Anat.*, 1936, *70*).

Ventricular surface of hypothalamus of human brain showing relative position and extent of some of the hypothalamic nuclei.

A = Anterior commissure; Hdm. = dorsomedial hypothalamic nucleus; Hp. = posterior hypothalamic nucleus; Hvm. = Ventromedial hypothalamic nucleus; Mm. = Medial mammillary nucleus; Mth. = Mammillo-thalamic tract (Bundle of Vicq d'Azyr); Pre. = preoptic nucleus; Pv. = paraventricular nucleus; So. = Supraoptic nucleus, (cf. Figs. 17, 408, 424, 427, 428, 429, 431.)

to respond less well to conditions of emergency. This impairment of functions is often not in evidence, and the animals may be able to run or fight as vigorously as normal controls. These results are important and cast grave doubt on the generally accepted view that the sympathetic-adrenal system is indispensable in mediating the adaptations necessary in times of stress (p. 731).

Higher Control of the Autonomic Nervous System.—1. CENTRES IN BRAIN STEM.—Superimposed over the connector cells of the autonomic system are "centres" in the *brain stem* controlling special parts of the system which are concerned with specific functions—*e.g.* the vasomotor (vascular tone) (p. 303), cardiac (heart rate) (p. 270), vomiting (p. 811), respiratory (p. 384), deglutition (p. 805), adrenaline-secreting (p. 730), and blood sugar regulating (p. 917) centres.

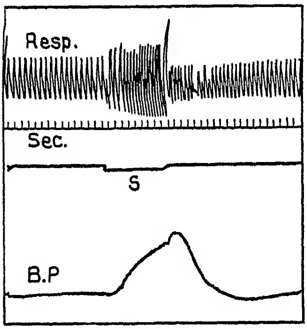


FIG. 449.—Effect of Hypothalamic Stimulation on Blood Pressure and Breathing. (Ranson, Kabat, and Magoun, *Arch. Neurol. Psychiat.*, Chicago, 1935, *33*, 468.)

Records from above downwards are: respiration, time in seconds, signal line, blood pressure. During fall of signal (S), region of lateral hypothalamic nucleus was stimulated. Note increase in rate and depth of breathing and rise of blood pressure.

2. **Rôle of the Hypothalamus.**—NUCLEI OF HYPOTHALAMUS. — The main nuclei in the hypothalamus may be grouped as follows (Fig. 448): (i) *Anterior*: (*a*) paraventricular; (*b*) supraoptic.

(ii) *Middle*: this group occupies the middle part of the tuber cinereum, and includes the nucleus tuberalis and the ventromedial, dorsomedial, and lateral hypothalamic nuclei.

(iii) *Posterior*: (*a*) posterior hypothalamic nucleus; (*b*) mammillary body.

CONNECTIONS AND FUNCTIONS.—These are summarized below.

(i) Efferent fibres from the hypothalamus pass from the middle and posterior nuclei to the brain stem "centres" mentioned in 1, *supra*, and thus *control the activity of the whole autonomic system* (sympathetic and parasympathetic).

(*a*) Stimulation of the *posterior* nuclei increases *sympathetic* activity; *e.g.* it produces a rise of blood pressure, secretion of adrenaline, quickening of the heart, and development of premature contractions (Fig. 449).

(*b*) Stimulation of the *middle* nuclei produces *parasympathetic* overaction, *e.g.* cardiac slowing, increased gastric secretion, blood flow and motility. Irritative lesions of the middle nuclei by stimulating the vagus, may produce hæmorrhagic erosions of the mucosa of the œsophagus, stomach or duodenum which may result in fatal perforation (p. 788).

(ii) Descending fibres from the frontal lobe (areas 6 and 4) (p. 629) and the globus pallidus (p. 656) end in the hypothalamus; the latter in turn gives rise to descending fibres which relay in the reticular nuclei, and are continued in the reticulospinal tracts to the spinal motor neurones (Fig. 408). The hypothalamus thus constitutes part of the *extra-pyramidal facilitatory pathway*. Lesions of the hypothalamus produce (experimentally) poverty of movement and catalepsy (p. 659) and may be responsible for some of the signs of Parkinson's disease (p. 660).

(iii) Fibres pass from the anterior nuclei to the *neurohypophysis* (posterior lobe of pituitary) to control the secretion of the antidiuretic hormone (ADH) (p. 47) and the oxytocic hormone (p. 1091). Though the hypothalamus does not directly innervate the *anterior pituitary* it forms chemical transmitters which regulate the activity of this gland (p. 931).

By reason of the connections set out in i–iii above, the hypothalamus is an important centre for *emotional exteriorization* (p. 665). Lesions which isolate the hypothalamus from the higher levels produce the rage reaction (pp. 664, 671). Lesions of the hypothalamus or its descending tracts may give rise to *emotional palsy* (p. 668). The hypothalamus is the principal centre for *temperature regulation* [1] (p. 473).

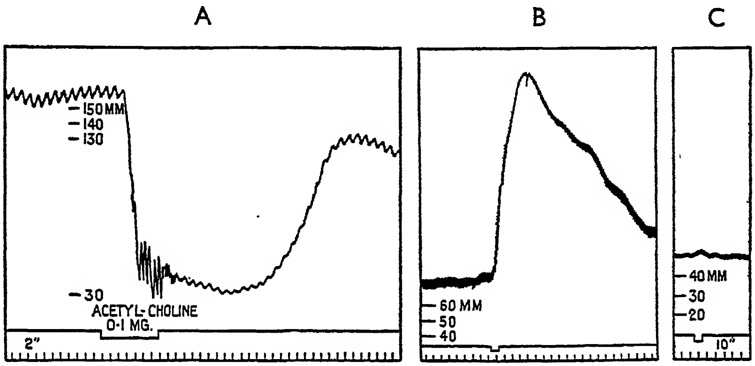


FIG. 450.—Action of Acetylcholine on Circulation. (Dale, *J. Pharm. exp. Therap.*, 1914.)

Records from above downwards: blood pressure, signal, time in 2 seconds (A, B), and 10 seconds (C)

A. Inject 0·1 mg. of acetylcholine. Marked fall of blood pressure and intense transient slowing of the heart (muscarine action). Between A and B inject 1 mg. of atropine. B. Inject 5 mg. of acetylcholine. Large rise of blood pressure (ganglionic or nicotine action). Between B and C inject large dose of nicotine (30 mg.) to paralyse autonomic ganglia. C. Repeat injection of 5 mg. of acetylcholine. No effect on blood pressure.

(iv) It is closely linked by many afferent and efferent fibres with the *thalamus and prefrontal cortex* (Fig. 428, p. 669). It is thus part of a nervous complex related to emotional states (p. 666), personality, and social behaviour (p. 670).

(v) It receives afferents from the hippocampal region via the fornix and discharges via the anterior nucleus of the thalamus to the cingular gyrus. The rôle of this connection is obscure.

Results of Lesions of the Hypothalamus.—Some of the syndromes observed clinically are briefly summarized below:

(i) Diabetes insipidus (p. 49).

(ii) Disturbed temperature regulation (p. 473).[1]

[1] The following points may be mentioned here: (1) Destruction of the hypothalamus may lead to prolonged hypothermia or hyperthermia. (ii) Localized heating or cooling of the hypothalamus leads to appropriate reactions which increase heat loss or heat production respectively, showing that the nerve cells are specifically sensitive to minute alterations in their temperature. (iii) The antipyretic group of drugs and the barbiturates (which render an animal poikilothermic) may produce their effects by an action on the hypothalamus.

(iii) Disturbances in carbohydrate and fat metabolism and in sexual function, resulting from anterior pituitary dysfunction (pp. 931, 936).

(iv) The occurrence of somnolence,[1] changes in muscle tone and impaired emotional exteriorization.

(v) Changes in personality presumably owing to deranged activity of the associated prefrontal cortex (p. 674).

(vi) Hypothalamic lesions may set up abnormal hunger which is satisfied by an excessive intake of food which in its turn produces *obesity*.[1] It will be recalled that prefrontal leucotomy is often followed, temporarily, by a ravenous appetite (p. 674).

3. Rôle of Cerebral Cortex.—The control of visceral activities by the cerebral cortex is considered on p. 671.

Mode of Action of Autonomic Nerves.—Chemical Transmission of Nervous Impulse.[2]—The general question of transmission of the nerve impulse was considered on p. 507. Before discussing this problem further in relation to the autonomic nervous system, we must be familiar with the peripheral actions of adrenaline (see p. 724), and of acetylcholine.

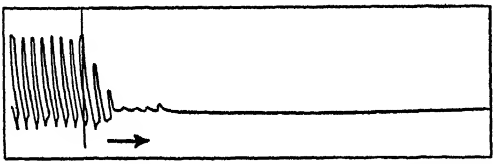


Fig. 451.—Action of Acetylcholine on Perfused Frog's Heart. (Dale, *J. Pharm. exp. Therap.*, 1914.)

Record of contraction of perfused frog's heart. At the vertical line the perfusing fluid was changed from pure Ringer's solution to one containing 1×10^{-8} of acetylcholine. Note arrest of the heart which follows an initial decrease in the force and rate of contraction.

Action of Acetylcholine.[3]—The peripheral *autonomic* effects of this drug may be grouped under three headings:

(i) *Direct* action: A peripheral dilator action on certain blood vessels.

(ii) "*Muscarine*" action: Acetylcholine stimulates all parasympathetically innervated structures by a peripheral action.

(iii) *Nicotine* action: Acetylcholine (like *small* doses of nicotine) stimulates all autonomic ganglia.

Action on the Circulation.—(i) If acetylcholine is injected intravenously in doses of 1×10^{-5} mg. or less (*e.g.* in the cat), it lowers the blood pressure. This effect is due to a direct dilator action on the walls of certain peripheral blood vessels as can be proved by perfusion experiments, *e.g.* on the blood vessels of the ear. This dilator action is abolished by atropine.

(ii) When larger doses (0·01–0·1 mg.) are injected a considerable fall of blood pressure results which is associated with, and is largely due to, a vagus effect on the heart, *i.e.* slowing and decreased contraction (Fig. 450, A). The cardiac effects are still obtained after section of the vagi, but are abolished by atropine, *i.e.* acetylcholine acts on the heart in the region of the vagal terminals.

(iii) After injection of atropine, small doses of acetylcholine no longer depress the heart or the blood pressure. Very large doses, however, *e.g.* 5 mg., produce a marked *rise* of blood pressure and cardiac acceleration, especially after previous destruction of the medulla and spinal cord (Fig. 450, B).

[1] Cf. footnote 1, p. 941.

[2] Dale, *Brit. med. J.*, 934, i, 1835. Loewi, *Proc. roy. Soc. B.*, 1935, *118*, 299.

[3] Dale, *J. Pharm. exp. Therap.*, 1914, *6*, 147; Symposium on Acetylcholine, *Bull. Johns Hopk. Hosp.*, 1948, *83*, 463.

The response is due to stimulation of the sympathetic ganglia and thus of the constrictor fibres to the blood vessels and the cardio-accelerator fibres; a discharge of adrenaline also occurs and is a minor contributory factor.

(iv) Following the injection of very large doses of nicotine which paralyse the autonomic ganglia (p. 715), the pressor effect of acetylcholine is completely abolished (Fig. 450, C), confirming the interpretation given above of its mode of action.

(v) Acetylcholine weakens, slows, or arrests the isolated perfused heart in concentrations of 1×10^{-8} or less (Fig. 451).

Action on Other Viscera.—There is no need to detail the actions of acetylcholine on the other viscera as it accurately repeats the effects of parasympathetic stimulation; thus it produces a secretion of tears, saliva, and gastric and pancreatic juice; it increases the movements of the œsophagus, stomach, small or large intestine (Fig. 452), and bladder; there is pelvic vasodilatation and erection of the penis. If injected directly (1 drop of 0·1% solution) into the pad of the cat's foot there is local vasodilatation and secretion of sweat (cf. p. 481). The uterus, though innervated by the sympathetic only, is stimulated by acetylcholine acting *directly* on the muscle coat.

Acetylcholine also has important actions on *skeletal muscle* (p. 514), the *central nervous system* (p. 530), and the *chemoreceptors* (p. 785). If applied directly to the *supraoptic nucleus* it stimulates it to discharge impulses to the posterior pituitary (p. 55).

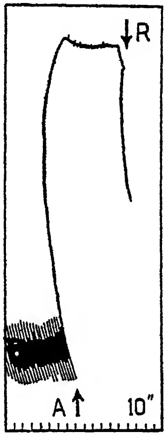


Fig. 452.—Action of Acetylcholine on Intestine. (Dale, *J. Pharm. exp. Therap.*, 1914.)
Record of movements of strip of intestine suspended in bath of warm oxygenated Ringer-Locke's solution. At A, acetylcholine was added to the bath. Note intense contraction of intestine. At R, fresh Ringer-Locke's solution was introduced into the bath and intestinal relaxation occurred.

The transient nature of the action of injected acetylcholine is due to its instability; it is rapidly hydrolysed in alkaline solution even at room temperatures to form choline which is about 100,000 times less potent than its acetyl derivatives. In blood and in the tissues generally there are varying concentrations of an enzyme, *cholinesterase* (p. 509), which greatly accelerates the rate of destruction.

Acetylcholine as Parasympathetic Chemical Transmitter.[1]—Loewi showed that when the vagus supply of an isolated perfused frog's heart is stimulated, a substance appears in the lumen of the ventricle which can be transferred to and produces a vagus effect on another ventricle (Fig. 453). There is good evidence that the substance is acetylcholine. The nervous impulses (spike potentials) which pass down the vagus, therefore, do not act directly on the heart muscle, but through a chemical intermediary or *transmitter* which is released at the vagal terminals (p. 508).

Not only the vagus, but all the parasympathetic nerves, both in amphibia and in mammals, produce their effects in exactly the same way. The study of the subject has been facilitated by the elaboration of biological methods which enable acetylcholine to be demonstrated in minute amounts and its

[1] Loewi, *Pflügers Archiv.*, 1921, *189*, 239. Brown, *Physiol. Rev.*, 1937, *17*, 485.

concentration determined with a fair degree of accuracy. The blood or perfusing fluid is collected before and during parasympathetic stimulation from the organ under investigation in an animal under the influence of eserine (this drug is used to preserve any acetylcholine which may be formed). A series of tests are then carried out, based on the following known actions of minute doses of acetylcholine :

(i) It lowers the blood pressure on intravenous injection; this fall is abolished by atropine.

(ii) It produces contraction of the back muscle of the leech (Fig. 454) or the rectus muscle of the frog previously treated with eserine.

(iii) It inhibits the frog's heart (Fig. 454) or the rabbit's auricle.

(iv) It is inactivated by treatment with alkali, but is unaffected by acid.

It the fluid under examination gives positive results with this series of tests and if in addition its relative activity on the different test objects is the same as that of acetylcholine there can be no reasonable doubt that it

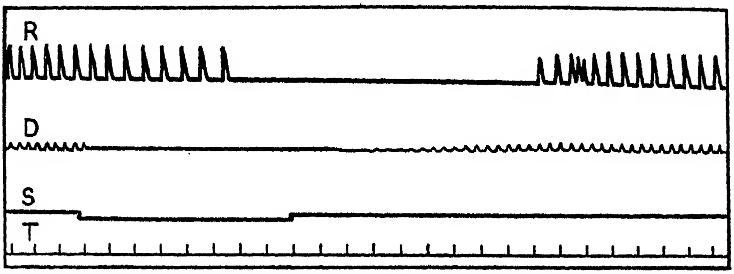


Fig. 453.—Liberation of Parasympathetic Chemical Transmitter in Frog's Heart. (Bain, *Quart. J. exp. Physiol.*, 1932.)

R and D represent the contraction of two isolated frogs' hearts. D (donor) is perfused with fluid which is then passed into the lumen of R (recipient). During the period indicated by the descent of the signal line S, the vagus nerve supply of D is stimulated, causing cardiac arrest. After a latent period the recipient heart (R) also ceases to beat. T = time in seconds.

contains acetylcholine. By matching the physiological results obtained with those produced by known concentrations of acetylcholine, one can determine the probable acetylcholine content of the sample (Fig. 454).

Acetylcholine is constantly being released at parasympathetic nerve ends under normal conditions because of parasympathetic "tone," *i.e.* the steady stream of impulses which are discharged at a low frequency from the nerve centres. The amount released is considerably increased by parasympathetic nerve stimulation. Acetylcholine is found in extracts of iris and of the ciliary bodies; the amount present is increased by third nerve stimulation; it is very important to note that it disappears completely after third nerve section and degeneration.

Action of Atropine.—Atropine does not interfere with the release of acetylcholine at parasympathetic postganglionic terminals; but it prevents the acetylcholine which is released from acting on the tissue cells. Subsequent parasympathetic stimulation consequently produces no functional changes in the organs supplied. Atropine is often said to "paralyse parasympathetic nerve endings"; this is not true; the current expression is that atropine "annuls the effects" of parasympathetic stimulation or "blocks" the parasympathetic nerve endings (cf. action of curare, pp. 516, 524).

Acetylcholine as Transmitter in Autonomic Ganglia.[1]—This question is fully considered on p. 523.

Chemical Transmitters of Antidromic Vasodilators in Dorsal Nerve Roots.—Stimulation of the peripheral end of a cut dorsal nerve root, *i.e.* towards the body surface, produces dilatation of the blood vessels (arterioles, capillaries, and venules) in the corresponding region of the skin and of the blood vessels of muscles. A limb enclosed in a plethysmograph consequently shows an increase in volume (Fig. 455). The fibres concerned have their cell bodies in the dorsal root ganglion, and not in the spinal cord: if the dorsal nerve root is cut peripheral to the ganglion (Fig. 195, point 2) and time is allowed to elapse for the appropriate nerve fibres to degenerate, this vasodilator effect is abolished; if the section is central to the ganglion (Fig. 195, point 1), the vasodilator fibres survive indefinitely. The nervous impulses in these experiments are called antidromic because they are transmitted in the reverse to the normal direction. If the *ventral roots are cut and allowed to degenerate*, it is found that subsequent peripheral stimulation of the cut dorsal root gives rise to a slow contraction of the skeletal muscles which have been deprived of their motor supply (the *Sherrington phenomenon*). The dorsal root fibres responsible for all the effects described are fine medullated nerves, conducting slowly at a rate of about 1 metre per second.

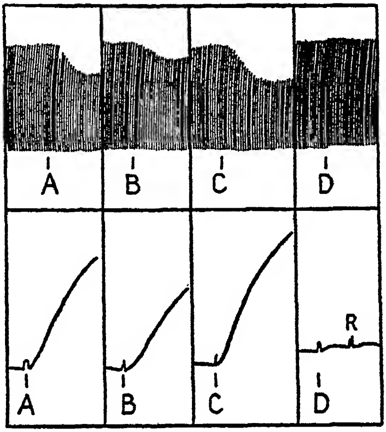


FIG. 454.—Estimation of Acetylcholine in Physiological Fluids. (Feldberg and Gaddum, *J. Physiol.*, 1934, *81*.)

Upper record, frog's heart; lower record, leech muscle treated with eserine.

A. Effects of fluid collected during nerve stimulation.

D. Control fluid.

B, C. Effects of acetylcholine 15 and 30 $\mu g.$ per litre respectively.

Note in A, B, C inhibition of heart and contraction of leech muscle. Test solution A is equivalent to about 20 $\mu g.$ of acetylcholine per litre.

The explanation of these remarkable results is as follows. Branches of the dorsal root fibres which reach the skin or muscles give off collaterals to supply the walls of the local blood vessels (Fig. 195). When the antidromic impulses reach these vessels they release chemical transmitters which are responsible for the effects described.

(1) *Effects on Muscle.*—There is evidence that acetylcholine is released at the ends of the dorsal root collaterals supplying the muscle vessels and is responsible for their dilatation. The acetylcholine thus released diffuses away and reaches the skeletal muscle fibres in the vicinity. It has been pointed out (p. 516) that motor denervation sensitizes tissues to the action of their normal transmitters. When skeletal muscle has been deprived of its motor nerve supply the sensitized tissue responds by contraction to the presence of acetylcholine in its vicinity. Intravenously injected acetylcholine likewise produces dilatation of muscle blood vessels and contraction of denervated skeletal muscle.

[1] Feldberg and Gaddum, *J. Physiol.*, 1934, *81*, 305. Brown, *Physiol. Rev.*, 1937, *17*, 485. Eccles, *Ergeb. Physiol.*, 1936, *38*, 339.

(2) *Effects on Skin.*—According to Lewis, the vasodilatation in the *skin* produced by peripheral dorsal root stimulation is due to the liberation there of a substance which resembles histamine (·H-substance) in its physiological properties. The question is further considered on pp. 309 and 323 *et seq.*

Evidence is presented on p. 513 that acetylcholine is the transmitter at motor end plates in *skeletal muscle.* Its relationship to transmission processes in the *central nervous system* is considered on pp. 530 *et seq.*

Chemical Transmitter of Sympathetic Postganglionic Fibres.[1]—The postganglionic fibres of the sympathetic produce their effects by releasing adrenaline or the closely related substance nor-adrenaline at their terminals. The transmitter will for convenience be referred to as adrenaline (p. 510). These results explain the so-called *sympathomimetic* action of adrenaline,

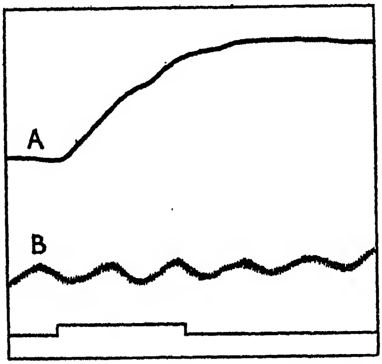


FIG. 455.—Vasodilators in Dorsal Nerve Roots. (Bayliss.)

Upper record (A) limb volume; lower record (B) blood pressure. Signal: stimulation of the *peripheral* end of a lumbar dorsal nerve root eight days after section between the spinal cord and the ganglion. Note the increase in limb volume (indicative of vasodilatation) without change in the level of the blood pressure.

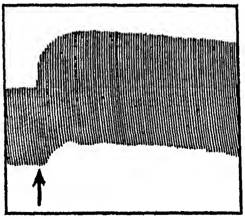


FIG. 456.—Release of Sympathetic Transmitter in Frog's Heart. (Loewi, *Pflügers Archiv.*, 1921, *189.*)

Record of movements of isolated perfused frog's ventricle. At the point marked by the arrow there was added to the perfusing fluid the contents of another frog's ventricle which had been subjected to sympathetic stimulation. Note as a result the increased amplitude of the contractions.

i.e. that it produces on injection the same results as sympathetic stimulation; no other result could be expected if the sympathetic itself produces its effects by releasing adrenaline at its terminals (cf. p. 724).

The following experiments may be quoted:

(1) Stimulation of the sympathetic nerve supply to the isolated frog's heart releases a substance in the ventricle which can be transferred to and causes sympathetic effects on another heart, *i.e.* increased force and rate of contraction (Fig. 456).

(2) Stimulation of the peripheral cut end of the sympathetic supply to structures in the hinder part of the body of the adrenalectomized cat, *e.g.* to the tail hairs, intestine, uterus, or bladder, causes acceleration of the denervated heart[2] (cf. p. 729). The maximal acceleration may occur several

[1] Bacq, *Ergeb. Physiol.*, 1935, *37*, 82; Rosenblueth, *Physiol. Rev.*, 1937, *17*, 514. Cannon and Rosenblueth, *Autonomic Neuro-effector System*, New York, 1937.

[2] Cannon *et al.*, *Amer. J. Physiol.*, 1931, *96*, 392; *97*, 365; 1932, *99*, 398.

minutes after nerve stimulation is discontinued; the effect is further delayed if the venous return from the stimulated region is temporarily arrested. Stimulation of the *peripheral* cut end of the sciatic nerve (which contains sympathetic fibres) causes adrenaline-like effects on distant organs, *e.g.* dilatation of the pupil (Fig. 457), inhibition of the intestine, a rise of blood sugar, and contraction of the spleen. All these results can be reproduced even when the postganglionic fibres to the responding organs have been cut and allowed to degenerate, and the adrenals removed. Chemical examination shows that the responsible exciting substance is adrenaline (or nor-adrenaline).

The adrenaline released at sympathetic postganglionic nerve endings exerts its effects preponderantly *locally* at the site of its liberation; small quantities, however, as has been shown, may escape into the general circulation and produce slight effects at a distance. As they release chemical intermediaries peripherally, all sympathetic (and parasympathetic) nerves might be called "secretory" nerves; or, alternately, one can regard the adrenal medulla as a large differentiated ganglion cell plus sympathetic ending. The essential difference between the sympathetic system and the adrenal medulla is that the former can act with greater precision on restricted parts of the body while the latter produces effects which are more generalized and diffuse.[1]

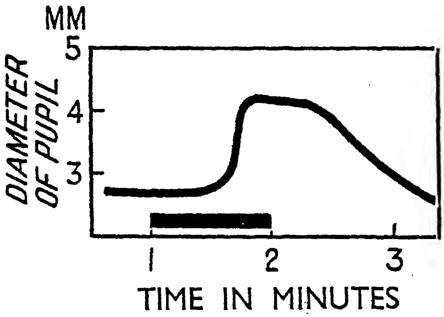


FIG. 457.—Release of Sympathetic Transmitter by Sympathetic Fibres to Limb. (Bacq, *Arch. internat. Physiol.*, 1933.)

Record of diameter of denervated pupil. During the period indicated by the signal, the peripheral end of the cut sciatic nerve was stimulated (this nerve contains postganglionic sympathetic fibres). After a latent period the pupil dilated.

THE ADRENAL MEDULLA[2]

The structure, comparative anatomy and embryology of the adrenal medulla are described on p. 943. The physiology of the medulla is dealt with here because its active principle (or principles) is as closely related functionally (as it is developmentally) to the sympathetic nervous system.

The *active principle* of the adrenal medulla has hitherto been thought to be *adrenaline,* which can be readily extracted from the medulla and can be demonstrated in the adrenal venous blood; recent careful work has shown however that the medulla also secretes the closely related substance called *nor-adrenaline* (which is adrenaline minus its terminal CH_3 group (p. 729)). The relative amounts of adrenaline and nor-adrenaline secreted vary a good deal for reasons that are unknown. There are significant differences in the

[1] The action of amine oxidase and ephedrine are considered on p. 510.
[2] Cannon, *Bodily Changes in Pain, Hunger, Fear and Rage,* 2nd edn., 1929.

physiological actions of adrenaline and nor-adrenaline. Until the matter has been further clarified, it is convenient to discuss the hormone of the adrenal medulla as though it were only adrenaline.

In the *resting condition* of the body it is not certain whether any secretion is discharged from the medulla, though there is some suggestive evidence. In various states of *emergency* or *stress* there is no doubt that adrenaline is poured into the circulating blood and acts mainly by reinforcing the activity of the sympathetic system.

Actions of Adrenaline.[1]—In general adrenaline affects all sympathetically innervated structures throughout the body and produces the same changes as does stimulation of their sympathetic nerve supply. (Adrenaline also acts on elements of the central nervous system.)

Briefly stated, the *effects* are (cf. pp. 708 *et seq.*):

(1) *Eye:* dilatation of the pupil (cf. p. 729) and retraction of the lids.

(2) *Heart:* the direct effects on the heart are acceleration, increased force of the beat, and increase in cardiac excitability and conductivity.

(3) *Blood Vessels:* constriction of the arterioles and capillaries in the skin; constriction of the splanchnic blood vessels; dilatation of the blood vessels supplying the heart and skeletal muscle.

(4) *Bronchi:* relaxation of the muscle coat.

(5) *Intestine:* inhibition of the intestinal wall and closure of the sphincters.

(6) *Liver:* conversion of glycogen into glucose. Contraction of the capsule of *spleen* and the wall of *gall-bladder.*

(7) *Bladder:* relaxation of the detrusor and contraction of the sphincter and trigone (p. 767).

(8) *Uterus:* varying effects are produced. In women the uterus is inhibited during labour and the puerperium.

(9) *Skin structures:* excitation of arrectores pili and other smooth muscle in the skin. In *man* no secretion of sweat is produced (but see p. 481).

(10) *Skeletal muscle:* it antagonizes muscular fatigue, enhances excitability, and antagonizes the action of curare (p. 521).

Certain points must be elaborated in greater detail.

1. **Circulation.**—A. Action in Animals.—(1) *Blood Pressure.*—Injection of adrenaline intravenously (in animals) produces a rapid and marked rise of arterial blood pressure (Fig. 458, A) (up to levels as high as 250 mm. Hg). The pressor effect is due to vasoconstriction, mainly in the skin and splanchnic area. The volume of a limb, kidney, or intestine (recorded with a plethysmograph) shrinks markedly, showing that the volume of blood in these parts is decreased because of contraction of the blood vessels. Similarly the blood flow through these regions descreases strikingly. The blood vessels of skeletal muscle dilate; thus, though the volume of an intact limb diminishes, the volume of a skinned limb increases after injection of adrenaline. The coronary vessels also dilate. The intensity of the action of adrenaline on different vessels is probably related directly to the richness of their sympathetic innervation. As might be anticipated, adrenaline has a relatively minor direct constrictor action on the cerebral and pulmonary vessels. The net result of

[1] The differences between the action of adrenaline and nor-adrenaline are mainly quantitative. Thus, if the dose of adrenaline necessary to produce a given effect is taken as 1, that of nor-adrenaline is: on cat's blood pressure, 0·8; rabbit ileum, 2; frog perfused heart, 33; rat non-pregnant uterus, 100.

these varying vascular effects is to *redistribute* the blood flow in a manner appropriate for conditions of stress. The flow is diverted from the skin and splanchnic areas and directed to the coronary and skeletal vessels which are actively dilated (*infra*). The rise of blood pressure further increases the rate of flow in this latter group of vessels; it also probably overcomes the slight constriction (directly produced) of the cerebral vessels, thus finally increasing the blood supply to the brain.

Adrenaline acts *peripherally* on the muscle coat of the blood vessels. It produces its pressor action after destruction of the medulla and spinal cord and after paralysis of the sympathetic ganglia by nicotine; the blood vessels of the skin are constricted even after section and degeneration of all the sympathetic nerves to the part.

(2) *Heart.*—Using the isolated mammalian heart perfused through the coronary vessels (p. 236) adrenaline produces an increase in the rate and force of cardiac contraction.

A pressor injection of adrenaline in the *intact* animal commonly slows the heart (Fig. 458, A). The rise of arterial blood pressure stimulates the presso-receptors in the carotid sinus and aortic arch; afferent impulses pass up to stimulate the cardio-inhibitory centre and reflexly, via the vagi, the heart rate is slowed (in other words, the *direct* accelerator action of adrenaline on the heart is reflexly overcome). This reflex cardiac slowing is abolished by section of the sino-aortic nerves, section of the vagi, or "blocking" of the vagal endings by atropine; it is also annulled when the blood pressure is prevented from rising by the use of a suitable compensator device (Fig. 458, B); under these experimental conditions adrenaline, intravenously injected, quickens the heart.

B. Action in Man.—An *intravenous* injection of adrenaline of *e.g.* 0·3 mg. produces a considerable rise of blood pressure (*e.g.* to 210 mm. Hg) associated with marked pallor owing to constriction of skin vessels. The skin temperature, as might be expected, falls especially in the hands and feet. The heart rate rises and there is increased excitability of the myocardium resulting in the development of numerous extrasystoles (p. 270); conduction in the bundle of His is speeded up.[1] If adrenaline is injected intravenously in *minute* amounts (1–2 μg.) there may be no change in arterial blood pressure or pulse rate or in the state of the skin vessels, but the muscle vessels in the calf or forearm dilate and the blood flow through them increases. These experiments prove conclusively that adrenaline *actively* dilates muscle vessels in man.

Subcutaneous injection of adrenaline (*e.g.* 0·2 c.c. of a 0·1% solution) causes constriction locally of arterioles and capillaries (p. 322), so that a cold pale patch of skin is produced.

If the heart has been arrested by an overdose of an anæsthetic, intra-cardiac injection of adrenaline may restore the rhythmic beat (cf. p. 294).

2. **Respiration.**—During the height of the rise of blood pressure which follows the intravenous injection of adrenaline in animals the respiratory movements often cease or become very shallow (Fig. 458, A). This *adrenaline apnœa*, as it is called, is mainly reflexly produced by the rise of blood pressure which stimulates afferent nerve endings in the aortic arch and carotid sinus

[1] Similar changes are observed when large doses of adrenaline are secreted by the cells of a tumour of the adrenal medulla (cf. p. 734).

(cf. Fig. 477). The apnœa is almost completely abolished after division of the vagi and denervation of the carotid sinuses, or if the rise of blood pressure is prevented from taking place by means of a compensating device (Fig. 458, B). In man, adrenaline apnœa does not occur; in fact in the "hypertensive

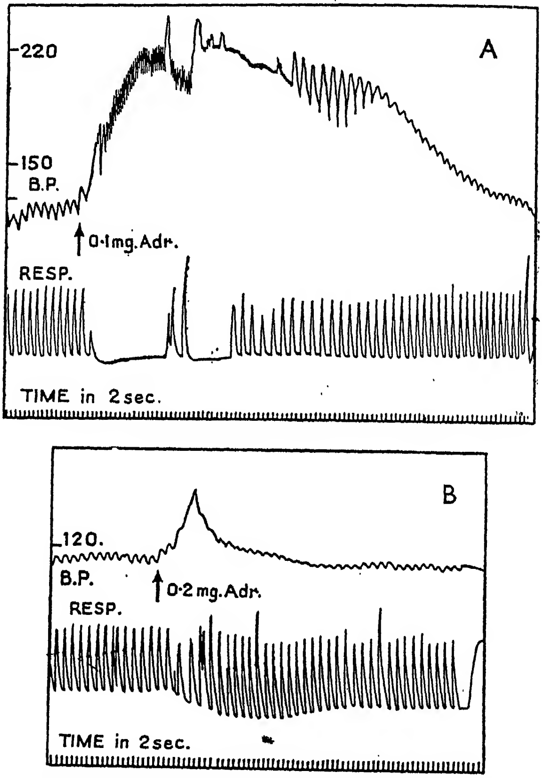


FIG. 458.—Action of Adrenaline on Blood Pressure and on Respiration. (Wright, *J. Physiol.*, 1930.)

Cat, chloralose anæsthesia. The records from above downwards represent blood pressure, respiration, and time in 2 seconds. The vagi are intact. In A, 0·1 mg. adrenaline was injected at the arrow. Note the marked rise of blood pressure and the obvious slowing of the heart rate. The breathing is inhibited at first and gradually returns to normal.

In B, twice the dose of adrenaline (0·2 mg.) was injected. By means of a special compensation device the rise of blood pressure was almost completely prevented. Note that there is no slowing of the heart, and only a transient slight decrease in respiration, following which the breathing becomes deeper.

This experiment demonstrates that the slowing of the heart and the respiratory depression which occur in the intact animal (A) are not due to the adrenaline itself but result (reflexly) from the rise of blood pressure.

crisis" of tumours of the adrenal medulla, the breathing becomes deep and rapid (p. 734).

Adrenaline increases the basal metabolism by about 20%, as the result of increased oxidation in the tissues generally. These changes come on within a few minutes of the injection and persist for about half an hour (p. 380).

The *bronchi* are relaxed; this action is the basis of the treatment of the spasm of bronchial *asthma* by subcutaneous injection of adrenaline (p. 409).

3. **Muscle.**—The actions on *skeletal muscle* are important. The muscles can work to better advantage owing to their improved blood supply; in addition (cf. p. 521) adrenaline increases the force of contraction both of normal and fatigued muscle in response to stimulation of its motor nerve. Adrenaline also has a marked *anti-curare action* (p. 522).

4. **Blood.**—(i) Adrenaline stimulates the conversion of glycogen in the liver into glucose which is then poured into the circulating blood. The

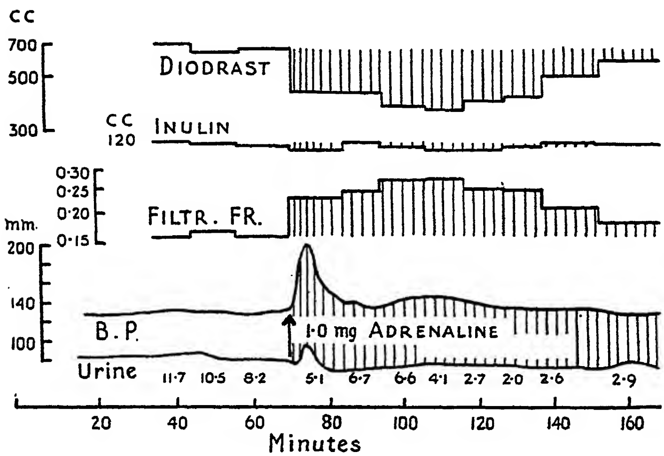


FIG. 459.—Effect of Adrenaline on Circulation and Kidney Function in Man. (Chasis *et al.*, *J. clin. Investig.*, 1938, *17*, 688.)

Diodrast=diodrast clearance=renal plasma flow per minute in c.c. Inulin=inulin clearance=glomerular filtration rate in c.c. per minute. Filtr. Fr.=filtration fraction=glomerular filtrate/renal plasma flow. B.P.=blood pressure; upper line=systolic, lower line=diastolic, pressure in mm. Hg. Urine=urine volume in c.c./10 minutes. Time in minutes.

At arrow, inject 1 mg. of adrenaline subcutaneously. Renal plasma flow decreases from 680 to 373 c.c./minute; glomerular filtrate is unchanged; filtration fraction rises from 16 to 30%. The volume of urine is reduced.

level of the blood sugar rises; if glomerular filtration of sugar is increased sufficiently, sugar is excreted in the urine (*glucosuria*). If the liver has previously been emptied of glycogen, a rise of blood sugar may still be produced by adrenaline either (*a*) because of stimulation of neoglucogenesis, or (*b*) owing to the conversion of muscle glycogen to lactic acid which enters the blood and is converted by the liver into blood glucose.

(ii) Blood coagulation time may be diminished.

(iii) Subcutaneous injection of adrenaline (*e.g.* 0·8 mg.) in man usually increases the red cell count, the hæmatocrit value (*e.g.* from 44·7 to 46·1%), the hæmoglobin concentration (*e.g.* from 14·5 to 15·2 g–%), and the plasma protein concentration (*e.g.* from 6·6 to 7·0 g–%). The rise in red cell count and the related changes have been attributed to mobilization of red cells

from depots, especially the spleen. Changes of the same magnitude, however, take place in splenectomized patients.[1] As the plasma protein concentration rises to the same extent as the red cells, the changes may be the result of simple hæmoconcentration due to movement of fluid out of the blood.

(iv) By causing a discharge of adrenal corticoids (*infra*) injection of adrenaline (*e.g.* 0·2 mg.) temporarily decreases the number of circulating *eosinophil leucocytes* (p. 951). The neutrophils increase in number; the lymphocyte count first rises and then falls.

5. **Kidney.**—The effects of pressor doses of adrenaline (1 mg. subcutaneously) in man are well summarized in Fig. 459. The blood flow through the kidney is cut down almost to half its original value owing to constriction of the renal arterioles. In spite of the diminished total renal blood flow the volume of glomerular filtrate formed is unchanged; the filtration fraction (*i.e.* the ratio of volume of glomerular filtrate/plasma flow) rises, *e.g.* from the normal 16% to 30%, indicating that glomerular capillary pressure has been substantially increased. The evidence as a whole proves that adrenaline constricts specifically the glomerular *efferent* arterioles (cf. p. 25). The *volume* of urine is decreased considerably; this change must be attributed to more complete reabsorption of water from the lumen of the tubules into the blood.

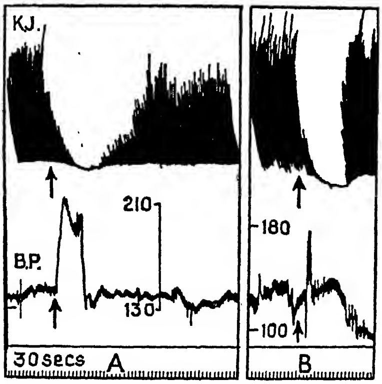


FIG. 460.—Action of Adrenaline on Spinal Cord. (Schweitzer and Wright, *J. Physiol.*, 1937.)

Cat.: Records from above downwards are: knee-jerk (elicited every 5 seconds), blood pressure, and time in 30 seconds. Between A and B *cut sino-aortic nerves*. At each arrow inject 0·2 mg. of adrenaline: inhibition of the knee-jerk and decreased quadriceps tone occurs in each case.

6. **Anterior Pituitary and Adrenal Cortex.**—Adrenaline stimulates the anterior pituitary causing the release of pituitary corticotrophin (ACTH); as a result there is increased secretion of adrenal corticoids (p. 950).

7. **Spinal Cord.**—Adrenaline acts on the central nervous system as well as peripherally. Thus in large doses it diminishes muscle tone and somatic reflexes (like the knee-jerk) by a direct depressant action on the spinal cord (there may be an initial excitatory action). These effects are independent of the associated changes in blood pressure or respiration, and are not due to constriction of the vessels supplying the cord or to afferent impulses in the sino-aortic nerves (Fig. 460).

Adrenaline may also act on the supraoptic nucleus stimulating the release of ADH from the neurohypophysis.

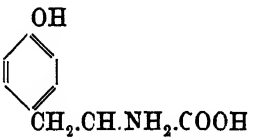
Site of Action of Adrenaline.—There is no doubt that the site of action of adrenaline on viscera and on skeletal muscle is a *peripheral* one; thus it acts on the isolated heart, the isolated intestine, or the blood vessels of the denervated limb. Its site of action is not the anatomical sympathetic nerve endings as it can still produce its effects on viscera after these nerve endings have degenerated. In fact the loss of the sympathetic nerve ends makes glands and smooth muscle respond in a more sensitive manner to

[1] Ebert and Stead, *Amer. J. med. Sci.*, 1941, *201*, 659.

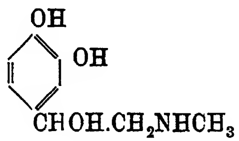
adrenaline, *e.g.* adrenaline introduced into the conjunctival sac dilates the pupil after extirpation of the superior cervical ganglion, but not when the nerve supply is intact (cf. p. 516). Skin blood vessels similarly respond more readily to adrenaline in man after, say, stellate ganglionectomy (cf. p. 359). Adrenaline thus acts directly on the muscle or gland cell *distal* to the nerve endings, presumably at the point where the chemical transmitter is released by sympathetic nerve impulses (p. 508, 722).

The action of adrenaline is intensified ("potentiated") by administration of ephedrine (p. 510) cocaine or thyroid. After injection of large doses of ergotoxin, doses of adrenaline which were previously pressor, produce a fall of blood pressure.

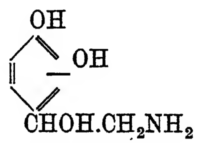
STRUCTURE OF ADRENALINE.[1]—Adrenaline and nor-adrenaline are related to the amino-acid tyrosine as shown by the formulæ below :



Tyrosine
para-hydroxy-phenyl-amino-propionic acid.



Adrenaline



Nor-adrenaline.

Injected adrenaline rapidly disappears from the blood; it is partly destroyed by the enzyme *amine oxidase* which is present both in blood and in tissues.

Functions of the Adrenal Medulla.—So far we have discussed the action of adrenaline from the pharmacological standpoint, as we would have considered that of any other drug. We must now consider a totally different question : What does the adrenal medulla do in the body ? The *methods of study* employed will first be described.

(1) HISTOLOGICAL.—When the adrenal is fixed with osmic acid, black granules can be seen in the medulla, which represent the adrenaline store of the gland. The changes produced in the adrenaline content under various conditions can thus be studied. It is found, for example, that exposure of a small animal to cold causes the granules to disappear; this observation suggests that cold stimulates the gland to secrete adrenaline to the point of exhaustion.

(2) DENERVATED HEART.[2]—This preparation is sensitive to adrenaline in minute concentrations, *e.g.* as low as $2{\cdot}5 \times 10^{-9}$. If any change in the state of such an animal produces cardiac acceleration it can be attributed either to adrenaline secreted by the medulla or to sympathetic transmitter which has diffused into the general circulation. If the acceleration is abolished by section of the nerves to the adrenals it must have been previously due to adrenaline secretion by the gland.

(3) CROSS CIRCULATION EXPERIMENTS (Fig. 461).—The venous blood

[1] Blaschko, in *The Hormones* (ed. Pincus and Thimann), N.Y., 1950, *2*, 601.

[2] The stellate and other ganglia supplying sympathetic fibres to the heart are extirpated and the cardiac fibres of the vagus are cut.

from the adrenal of a dog (B) is drained into the circulation of an adrenalectomized dog (C), which serves as the test preparation. The blood pressure and volume of an organ (*e.g.* spleen or intestine) in C are recorded. A rise of pressure and constriction of the spleen or intestine of C are taken to indicate increased secretion of adrenaline in B; a fall of blood pressure and dilatation of the viscera of C indicate inhibition of adrenaline secretion in B. By these means it has been shown, for example, that a *fall* of blood pressure in B *stimulates* adrenaline secretion, while a *rise* of blood pressure *inhibits*

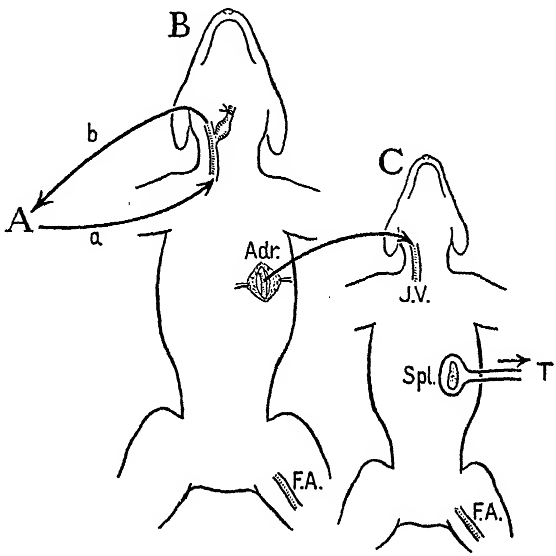


FIG. 461.—Diagram to illustrate Cross Circulation Technique. (Tournade and Heymans.)

Animal A (not drawn) perfuses the isolated carotid sinus region of B; *a* brings blood from the carotid artery of A and *b* leads the blood back to the jugular vein of A. The carotid sinus of B is excluded from the circulation but is connected by means of its afferent nerve with the medulla. Blood is drained from the adrenal vein of B into the jugular vein of the adrenalectomized animal C. The blood pressure of B and C are recorded from the femoral artery. The volume changes of the spleen of C are recorded by means of a plethysmograph connected with a tambour.

secretion. Rigid proof can also be obtained in this manner that the splanchnics are the secretory nerves to the adrenal medulla (Fig. 462).

Nervous Control of Adrenaline Secretion.—The splanchnic fibres which end in the adrenal medulla are *preganglionic*; they end directly round the medullary cells themselves. As might be expected, the chemical transmitter between these preganglionic fibres and the adrenal cells is acetylcholine (*i.e.* as in autonomic ganglia generally) (cf. p. 523). When the peripheral end of the splanchnic nerve is stimulated, adrenaline is secreted; this can be demonstrated by leading the blood from the adrenal vein into another animal (the recipient), where it produces a rise of blood pressure and vasoconstriction (Fig. 462). There is probably a *centre* in the brain stem which exercises a higher control. The activity of this centre may

be modified by afferent impulses, especially along the sinus and aortic nerves. The secretion of adrenaline ceases after bilateral splanchnic nerve section.

Regulation of Adrenaline Secretion.—CONTINUOUS SECRETION.—It has hitherto been generally accepted that adrenaline is not secreted under resting conditions. The following experiments are offered in support of the contrary view: Adrenalectomy in the dog causes the blood pressure to fall and the rate of the denervated heart to decrease (in acute experiments). It is possible to drain the blood from the adrenal of another animal—the donor B—into the jugular vein of the adrenalectomized recipient C. The entrance of the adrenal blood from the resting animal (B) causes the blood pressure and heart rate in C to rise and the spleen volume to decrease (indicative of arteriolar constriction); if the venous inflow from B is arrested the blood pressure in C falls and the spleen volume increases (Fig. 463). These observations suggest that a resting secretion of adrenaline takes place in B. It can be fairly argued, however, that animal B can hardly be compared with a normal *intact* animal in view of the complicated operative procedure to which it has of necessity been subjected; the question of continuous secretion under *normal* resting conditions is still undecided.

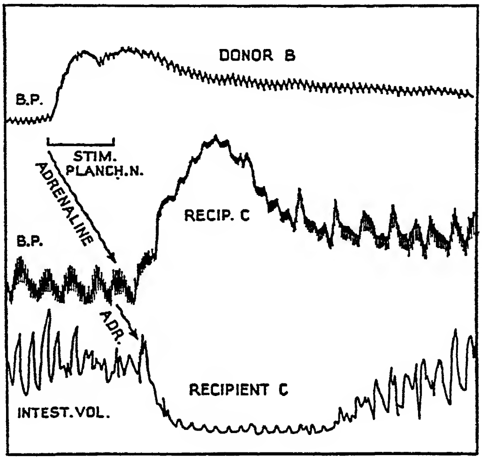


FIG. 462.—Splanchnic Control of Adrenaline Secretion. (Tournade.)

The animals B (donor) and C (recipient) are connected up like B and C in Fig. 461.
The records from above downwards are: Blood pressure of donor B, blood pressure of adrenalectomized recipient C, intestinal volume of recipient C.
During the period marked by the signal line the splanchnic nerve in B is stimulated. There is the usual rise of blood pressure in B. Adrenaline is secreted and enters the circulation of C and produces there a rise of blood pressure and constriction of the intestinal vessels.

Adrenaline Secretion during Conditions of Stress.—By the various methods described above it can be shown that adrenaline secretion is stimulated by (i) *physical exertion and certain emotional states*, (ii) *exposure to cold*, (iii) *fall of arterial blood pressure*, (iv) *asphyxia and cerebral anœmia*, (v) *anœsthesia*, (vi) *stimulation of afferent nerves*, (vii) *hypoglycœmia* (p. 915). The adrenaline thus poured out acts mainly on various sympathetically innervated structures and enables the body to deal more adequately with the state of emergency.[1] It also causes a secretion of adrenal *corticoids* which are thought to be of value in the general response to stress (p. 947). As the adrenal cortex is not innervated by the sympathetic, adrenaline in this instance is producing effects which cannot be directly brought about by the sympathetic.

[1] A sudden rise of arterial pressure in the "resting" experimental animal reflexly *decreases* adrenaline secretion.

presumed to perform the following functions: (i) aid the redistribution of blood in the body to the active muscles; (ii) increase the force and rate of the heart and the coronary blood flow, and thus enable the heart to cope better with the larger venous return; (iii) mobilize liver glycogen and provide glucose for the active tissues; (iv) diminish fatigue in skeletal muscle (p. 521); (v) in some species (but probably not in man) increase the oxygen-carrying power of the blood by discharging the red cells stored in the spleen; (vi) relaxation of the bronchi (facilitating ventilation of the alveoli), dilation of the pupil, and closure of the sphincters, may be other effects of value to the organism. But as already explained the response to such states of stress is little impaired in animals after exclusion of the adrenal medulla and even after complete sympathectomy.

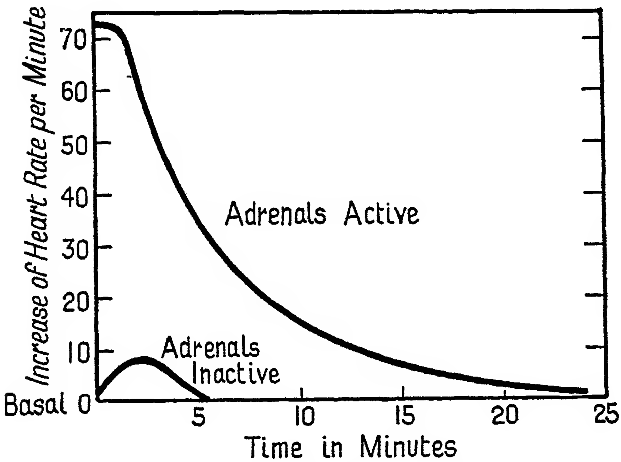


FIG. 464.—Adrenaline Secretion as a Result of Emotion. (Cannon and Britton, *Amer. J Physiol.*, 1927.)

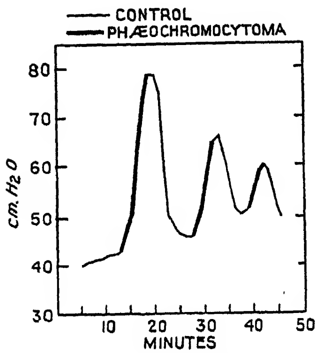
Denervated heart preparation. Cat excited by barking dog for 2 minutes, then removed from cage and rested quietly on a cushion. With adrenals inactivated by denervation, there was only a slight increase in heart rate, which rapidly subsided. With active adrenals there was a faster heart rate, which persisted for over 20 minutes though the cat was excited for only 1 minute.

EXPOSURE TO COLD (cf. p. 480).—When a cat is exposed to a current of cold air or if cold water is introduced into the stomach, the rate of the denervated heart *increases* by 10–33%; if the experiment is repeated after inactivation of the adrenals the rate *diminishes* because of the direct effect of cold on the heart. Clearly, adrenaline is secreted under the circumstances of the experiment and helps to maintain body temperature; it presumably *stimulates cellular metabolism* and increases heat formation, constricts the skin vessels and so diminishes heat loss, and mobilizes liver glycogen to provide energy. Adrenaline secretion is the primary delicate reaction to cold; shivering is the second grosser compensatory mechanism. Cannon expresses it thus: from the standpoint of temperature regulation, adrenaline secretion is a "fine" adjustment, shivering a "coarse" adjustment.[1]

[1] A bout of shivering may give rise to a sudden intense increase of the metabolism, *e.g.* to 60 or 80% above normal.

ADRENAL AND BLOOD PRESSURE REGULATION.—Adrenaline secretion is adjusted to assist in steadily maintaining the normal resting blood pressure. Thus, if the blood pressure is lowered by hæmorrhage or by stimulation of the peripheral end of the vagus, adrenaline secretion is increased; if the blood pressure is raised (for example, by injection of a large volume of blood) the secretion is arrested. Even very slight changes in the blood pressure may appropriately modify the rate of adrenaline secretion.

The secretion of adrenaline to help stabilize the blood pressure is reflexly regulated by the aortic and sinus nerves; neither the adrenal glands nor the bulbar centres controlling them are acted on directly by the level of the blood pressure.



G. 465.—Adrenaline in the Blood in Adrenal Medulla Tumour (Phæochromocytoma). (Prinzmetal *et al.*, *Ann. Surg.*, 1937, *16*.)

Plasma was collected during an attack of hypertension in which the blood pressure was 300 mm. systolic and 200 mm. diastolic. It was perfused through the blood vessels of a rabbit's ear during the period marked by the thickened line on the tracing, and the perfusion pressure (along the ordinate) recorded in cm. H_2O. A rise of perfusion pressure indicates vasoconstriction. Note the marked constriction that was obtained on each occasion; relaxation occurred when plasma from a normal control subject was substituted. The constrictor effect was abolished by ergotamine (as is the case with adrenaline). Following removal of the tumour the blood pressure fell to 130/90 and the plasma gave negative results. The tumour contained adrenaline.

At rest, the aortic and sinus nerves exert, in the main, a tonic inhibitory influence over adrenal activity; division of the sino-aortic nerves results in increased adrenaline secretion.

Adrenal Medulla Tumour (Phæochromocytoma).[1]—In this disease there is tumour growth of the cells of the adrenal medulla, which from time to time secrete large doses of "adrenaline."[2] The changes produced in an attack are in the main those that might be expected from the known actions of adrenaline. There is severe palpitation "as though the heart would burst through the ribs"; the pulse rate is rapid, *e.g.* 120 per minute; there is intense cutaneous vasoconstriction; the hands and face are cold and pale (sometimes blue); there may be profuse and drenching sweating; the blood pressure may rise up to 300 mm. Hg systolic and 200 mm. Hg diastolic ("hypertensive crisis"). The pupils are widely dilated and do not react to light. The breathing (unexpectedly) becomes deep and rapid and there may be headache, nausea, vomiting and considerable tremulousness. As the attack passes off flushing sets in and the blood pressure falls—to normal in early cases, but not in chronic ones. In the latter, progressive structural changes develop in the blood vessels especially those in the kidney (ultimately producing renal failure) and in the retina; the changes resemble those found in malignant hypertension (p. 354). Systemic venous blood collected during such an attack has been shown (by perfusion and other methods) to contain an *adrenaline-like* substance. In the case illustrated by Fig. 465 removal of the tumour restored the blood pressure to normal and the adrenaline disappeared from the blood.[1]

These tumours may contain both adrenaline and nor-adrenaline in large

[1] Wilkins *et al.*, *Arch. int. Med.*, 1950, *86*, 51. Calkins *et al.*, *J. clin. Endocrin.*, 1950, *10*, 1.
[2] "Adrenaline" = adrenaline + nor-adrenaline in varying proportions.

amounts, but the ratio between the two substances varies widely. Adrenaline and nor-adrenaline are excreted in corresponding proportions in the urine.[1]

VISCERAL AFFERENT NEURONES.[2]

The topics to be considered cover a wider field than is indicated by the title of the section. They include:

(i) Physiology of the vascular pressure-receptors and chemoreceptors.
(ii) Cutaneous hyperalgesia.
(iii) Referred pain, tenderness and rigidity from deep somatic structures.
(iv) Pain of cardiac ischæmia; mechanism of ischæmic pain.
(v) Sensibility of pleura, pericardium and peritoneum.
(vi) Sensibility of abdominal viscera.

The visceral afferents are chiefly medullated fibres of varying size with an admixture of unmyelinated fibres, all of which have their cell bodies in the dorsal root ganglia or their cranial homologues.

Afferent Nerves from the Aorta, Great Veins, and Heart.—(1) *Aortic ("Depressor") Nerve.*—Some fibres arise in the adventitia of the *arch of the aorta* (and the beginning of its *main branches*) in sensory endings which resemble those found in the carotid sinus (p. 737); some fibres may also come from the wall of the left ventricle. The aortic fibres reach the medulla oblongata mainly in the vagus trunk. The receptors in the aortic wall are sensitive to changes in the *blood pressure*, and send up impulses which reflexly regulate blood pressure, heart rate and rhythm, adrenaline secretion, and pulmonary ventilation (p. 738). Other fibres arise in the *aortic body* which lies near the aortic arch; it resembles the carotid body in structure and likewise functions as a *chemoreceptor* (p. 738), (Fig. 466).

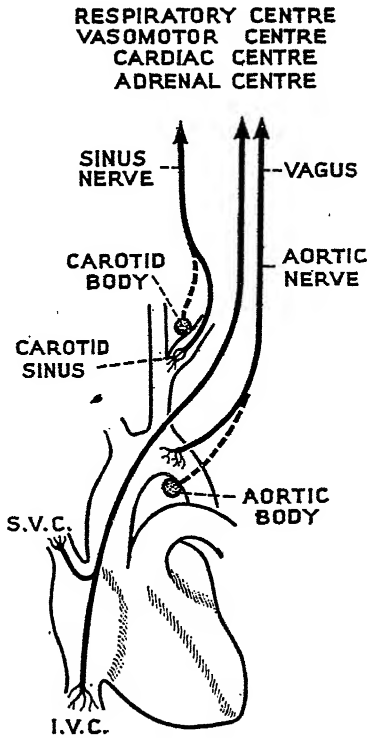


FIG. 466.—Afferent Nerves of Cardio-vascular System.

S.V.C., I.V.C.=superior and inferior vena cava. Note the afferent fibres arising from the carotid sinus and carotid body, the aortic arch, aortic body, and superior and inferior venæ cavæ.

(2) *Afferents from Great Veins.*—Afferents arise from the roots of the

[1] Engel and Euler, *Lancet*, ii, 1950, 387.

[2] Hurst, *Sensibility of Alimentary Canal*, London, 1911. Mackenzie, *Symptoms and their Interpretation*, 4th edn., London, 1920. Ranson, *Physiol. Rev.*, 1921, *1*, 477. Morley, *Abdominal Pain*, Edinburgh, 1931. Capps, *Clinical Study of Pain*, New York, 1932. Schweitzer, *Irradiation Autonomer Reflexe*, Basle, 1937. Lewis, *Pain*, New York, 1942. Symposium on Pain, *Res. Publ. Assoc. nerv. ment. Dis.*, 1942, *23*.

superior and inferior venæ cavæ and pulmonary veins (Fig. 467) (and perhaps from the right auricle), and pass in the vagus nerve to the medulla. These fibres reflexly adjust the rate of the heart to variations in the venous return (p. 272); they also reflexly affect respiration (p. 404).

(3) *Afferents from the Heart.*—Large afferent medullated fibres pass from the heart in the middle and inferior cervical branches of the sympathetic to the corresponding ganglia and thence to the upper two thoracic ganglia; other fibres go direct to the upper thoracic ganglia. The afferents then travel

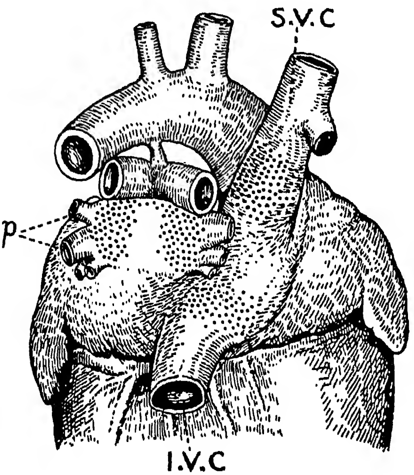


FIG. 467.—Distribution of Afferent (receptor) areas on Venous side of Circulation. (Nonidez, *Amer. J. Anat.*, 1937, *61*.)

Posterior surface of heart showing (stippled areas) extent of receptor areas on terminal portions of superior vena cava (S.V.C.), inferior vena cava (I.V.C.) and pulmonary veins (*p*). No receptors are found in the auricles themselves.

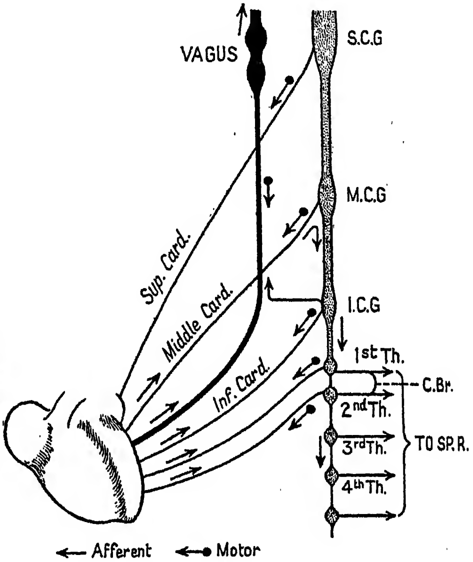


FIG. 468.—Afferent Nerve Supply of the Heart.

S.C.G., M.C.G., I.C.G. = Superior, middle, and inferior ganglia of cervical sympathetic; C.Br. = communicating branch between rami from 1st and 2nd thoracic ganglia; 1st, 2nd, 3rd, 4th Th. = white rami from thoracic sympathetic ganglia to corresponding spinal nerves to enter dorsal nerve roots (SP.R.); Sup. Card., Middle Card., Inf. Card. = Superior, middle, and inferior cardiac branches of the cervical sympathetic.

in the white rami to join the upper four (or more) thoracic dorsal nerve roots, and thus enter the spinal cord (Fig. 468). These afferents subserve *pain* sensibility. The referred pain in angina pectoris (p. 750) and coronary disease is felt in the arm and chest, chiefly in the distribution of the eighth cervical to the fourth thoracic dorsal roots. According to Leriche, the *anginal syndrome* may be reproduced in man by stimulation of the stellate ganglion; further, injection of procaine into the ganglion may arrest a spontaneous attack of angina.

Carotid Sinus.[1]—The carotid sinus is a dilatation normally present at the bifurcation of the common carotid artery; in man it is usually restricted

[1] Heymans, Bouckaert, and Regniers, *Le Sinus Carotidien*, 2nd edn., Paris, 1933.

to the first part of the internal carotid (Figs. 469, 470). Lying in the wall of the sinus between the white fibres of the adventitia are characteristic sensory nerve endings (Fig. 471). They have the typical structure of *stretch* receptors and resemble those found in the aortic arch and in tendons. Like the aortic endings, the sinus endings respond to alterations in the level of the blood pressure. The afferent fibres from the sinus pass mainly in the glossopharyngeal nerve to the medulla.

Fig. 469.—Anatomy of Carotid Sinus. (De Castro, *Trav. Lab. Rech. Biol.*, 1928.)

Bifurcation of common carotid artery, showing carotid sinus as dilatation on origin of internal carotid artery. + : position of sensory nerve endings. o >: regions rarely or slightly innervated.

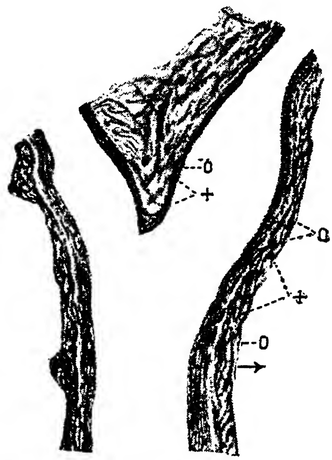


Fig. 470.—Section through Carotid Sinus Region. (De Castro, *Trav. Lab. Rech. Biol.*, 1928.)

Longitudinal section of bifurcation of common carotid artery. Note the carotid sinus—the dilatation at the commencement of the internal carotid artery. Points marked + are regions of maximum sensory innervation. The sensory nerve endings can be seen to lie deep in the adventitia. At points marked O nerve endings are rarely present.

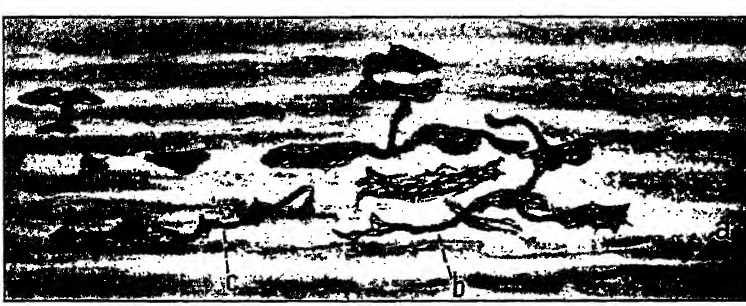


Fig. 471.—Sensory Nerve endings in Carotid Sinus. (De Castro, *Trav. Lab. Rech. Biol.*, 1928.)

Sensory terminals (b, c)lie between white fibres (a) of adventitia.

Carotid and Aortic Bodies.[1]—The *carotid body* lies at the carotid bifurcation near the carotid sinus; the *aortic body* lies in the thorax between the pulmonary artery and the ascending aorta (Fig. 466). Both bodies are richly supplied with blood from special arteries; the structure and functions of both are identical. They function as *chemoreceptors* sampling changes in the chemical composition of the blood. They are affected by alterations in CO_2 and O_2 tension, H^+ ion concentration, and the presence of certain drugs like nicotine, acetylcholine, and cyanide. The afferent impulses arising in these bodies reflexly modify respiration mainly (p. 745), and the circulation (vasomotor tone and heart rate) to a lesser degree.

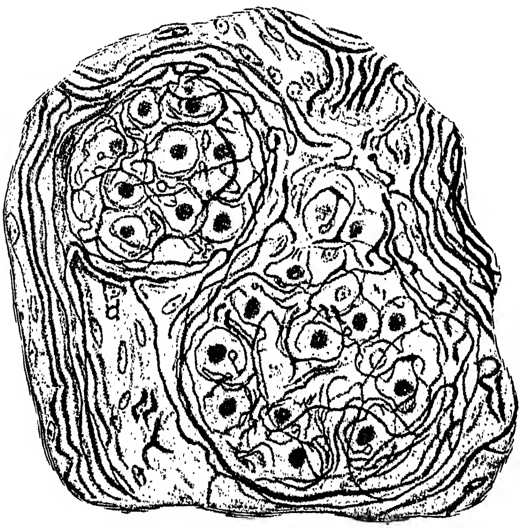


FIG. 472.—Structure of Carotid Body. (De Castro, *Trav. Lab. Rech. Biol.*, 1926.)

d—Medullated sensory fibres ending in intimate relationship with the epithelial cells of the carotid body.

The carotid body consists of masses of glandular-looking cells arranged in characteristic clumps (Fig. 472). The nerve supply is very rich; there are numerous afferent filaments which lie close to the gland cells and the internal lining of the numerous vascular sinuses. They are thus well placed to detect alterations in the composition of the blood (Fig. 473).[2]

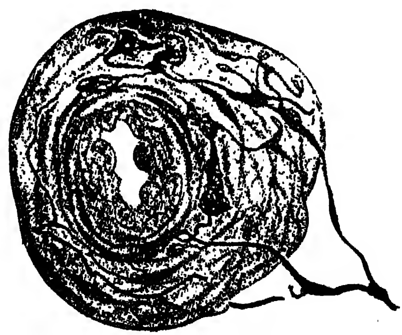


FIG. 473.—Innervation of Blood Vessels of Carotid Body. (De Castro, *Trav. Lab. Rech. Biol.*, 1926.)

Afferent nerve endings in the adventitia of a small blood vessel in the carotid body.

Functions of Carotid and Aortic Sinuses.—(1) *Results of Stimulation of Aortic or Sinus Nerves.*—Stimulation of the central end of the *aortic* nerve (or the central end of the vagus nerve, which contains aortic fibres) usually reflexly slows and weakens the heart and lowers the blood pressure (Fig. 474).

[1] De Castro, *Trav. Lab. Rech. Biol. Inst. Cajal*, 1926, *24*, 365.

[2] The large *peripheral arteries* have numerous sensory endings (some unencapsulated, others resembling end bulbs) in the adventitia, leading to medullated afferent fibres. Perfusion experiments on the isolated innervated hind limb show that these endings do not react to pressure changes but respond when irritant solutions are employed; they probably subserve *pain* sensibility.

(i) The slowing of the heart is mainly due to reflex stimulation of the vagus nucleus, increasing the discharge along the (efferent) vagal cardio-inhibitory fibres. After bilateral vagal section or injection of atropine, only slight reflex slowing is obtained (because of inhibition of sympathetic accelerator tone). The aortic reflex thus employs the antagonistic nerve supply of the heart *synergistically* to produce the desired effect, *i.e.* the vagal inhibitory fibres are stimulated and the sympathetic accelerator fibres are simultaneously depressed.

(ii) The fall of blood pressure is due (*a*) to decreased cardiac output, and (*b*) to inhibition of the normal tonic discharge of the vasomotor centre, leading to generalized vasodilatation and consequently to a decrease in peripheral resistance. When the cardiac effects are excluded by vagal section or atropine, central stimulation of the aortic nerve still produces a fall of blood pressure (Fig. 475), which is less marked because it is due solely to vasodilatation.[1] The secretion of *adrenaline* may also be arrested (cf. p. 734).

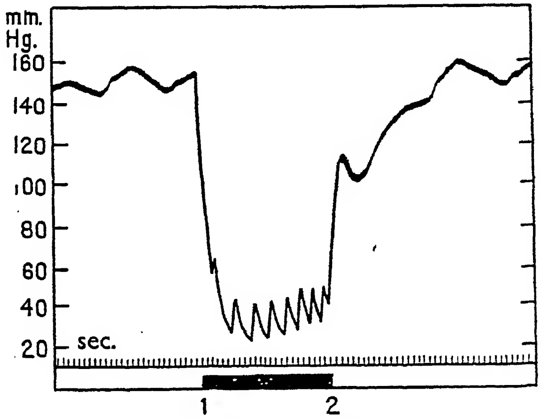


Fig. 474.—Reflex Effects on Blood Pressure and Heart Rate produced by Stimulation of Central End of Aortic Nerve in the Dog (between points 1 and 2). (Schweitzer, *Irradiation Autonomer Reflexe*, Basle, 1937.)

Note fall of blood pressure and extreme cardiac slowing. Time in seconds. (The individual heart beats cannot be seen during the control period owing to the high heart rate.)

The aortic reflex illustrates certain features of reflex inhibition in the autonomic system (particularly recruitment and after-discharge).[2] Increasing the frequency of stimulation applied to the central end of the vagus increases the *rate* and *extent* of the fall of blood pressure, and sometimes prolongs the duration of the recovery process (Fig. 475); similar effects are produced by increasing the *strength* of stimulation. Prolongation of the *duration* of stimulation increases the extent of the fall of blood pressure (cf. Fig. 340).

Sometimes central vagus or aortic nerve stimulation may produce a *pressor response* consisting of reflex cardiac acceleration, and vasoconstriction from *stimulation* of the vasomotor centre; the blood pressure rises. This

[1] For the methods employed for studying changes in the peripheral circulation, cf. pp. 304 *et seq.*
[2] Wright, *J. Physiol.*, 1928, *16*, 387.

experiment proves that there is an admixture of some pressor (excitatory afferent) fibres among the far more numerous depressor (inhibitory afferent) fibres in the aortic nerve (cf. p. 312). These pressor afferents come from *chemoreceptor* endings (and not from pressure-receptors) (p. 746).

Stimulation of the sinus nerve produces results identical with those described above for the aortic nerve.

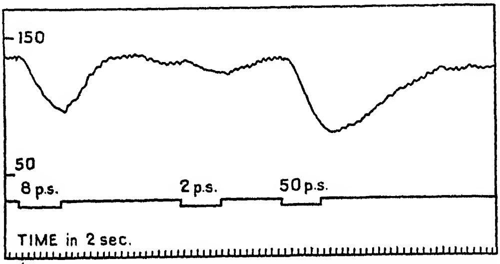


FIG. 475.—Effect on Aortic Depressor Reflex of Increasing Frequency of Stimulation. (Wright, *J. Physiol.*, 1928.)

Cat. Decerebrate. Right vagus cut. Stimulate left central vagus; stimuli are approximately of equal duration. 1st, 8 double shocks per second; 2nd, 2 d.s.p.s.; 3rd, 50 d.s.p.s. Time in 2 sec. The slight fall of blood pressure which precedes the second stimulus is due to slight pulling on the nerve during the adjustment of the electrodes.

(2) *Effects of Stimulating Nerve Endings in Carotid Sinus.*—(i) The sinus nerve endings in man can be stimulated by applying external digital pressure. Such compression produces the expected reflex cardiac slowing, vasodilatation, and fall of blood pressure. Very marked slowing may be obtained in arteriosclerotic subjects, probably because the sensory endings are pressed against the hardened media and so receive vigorous stimulation. Sometimes complete cardiac arrest may result (and syncopal symptoms develop), or various

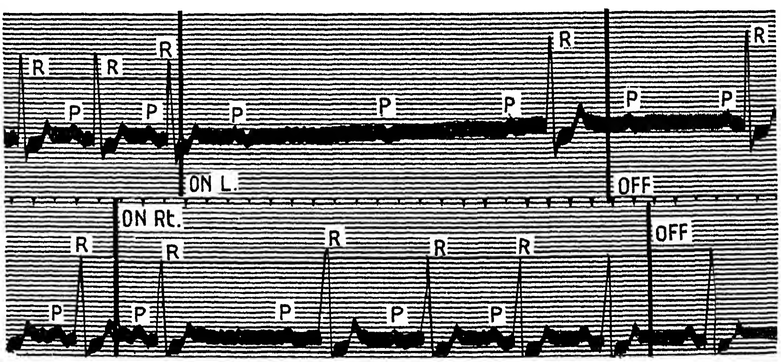


FIG. 476.—Effects of Carotid Sinus Compression on Human Heart. (Ward and Wright.)

Electrocardiographic records (Lead II.) in a man with marked thickening of the carotid arteries. During the period between signalling "on" and "off" the carotid sinus was compressed. (Upper record, left side; lower record, right side.)

In the upper record, note transient complete heart block and slowing of the auricular waves (P). In the lower record, note marked prolongation of the P–R interval indicating impaired conduction in the bundle of His. (Cf. Figs. 161, 162, pp. 283 *et seq.*, illustrating clinical heart block.)

grades of heart block may be produced (Fig. 476). Attacks of paroxysmal tachycardia (p. 290) may occasionally be arrested for a short time in this way.

In some patients the sinus nerve endings may be so *hyperexcitable* that the slightest pressure on the skin arrests the heart and produces loss of

consciousness and epileptiform seizures.[1] Denervation of the sinuses abolishes the attacks.

(ii) The carotid sinus endings can be stimulated under comparatively physiological conditions in suitably planned *perfusion* experiments. The technique is as follows: One cannula is introduced into the distal end of the cut common carotid artery and another into the central end of the external carotid artery (or its lingual branch). The internal carotid artery is tied off beyond the carotid sinus, as are all branches arising between the points of insertion of the cannulæ. The nerve supply of the sinus is preserved.

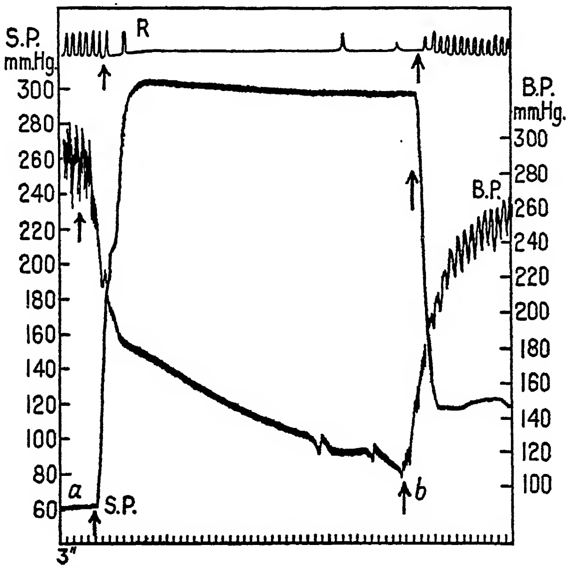


FIG. 477.—Reflex Effects of Rise of Carotid Sinus Pressure on Circulation and Respiration. (Heymans and Bouckaert, *J. Physiol.*)

R = Record of respiration; B.P. = Blood pressure; S.P. = Pressure in isolated innervated carotid sinus. During the period *a b* (indicated by arrows) the carotid sinus pressure was artificially raised. Respiration was reflexly inhibited and the systemic blood pressure was reflexly lowered. (Both vagus nerves had been cut.)

The pressure in this *carotid bifurcation preparation*, or *isolated carotid sinus preparation*, as it is more usually called, can be varied by perfusing it from a pump or from the circulation of another animal. If the latter technique is used the central end of the carotid artery of the donor (A) is connected with the distal end of the carotid of the recipient (B), and thus to the carotid sinus region. The blood is returned to A by anastomozing the second cannula in B with the central end of the jugular vein of A. The effects of varying carotid sinus pressure in B on various functions in the animal, *e.g.* heart rate, arteriolar tone, adrenaline secretion, respiration, action potentials in the sinus nerve, skeletal muscle tone, and general visceral activity are carefully studied.

(3) *Effects of Changes in Sinus Pressure.*—(i) A rise in sinus pressure

[1] The vagus nerve trunk in man (unlike the carotid sinus endings) is comparatively inexcitable to mechanical stimulation.

produces reflexly: (*a*) slowing of the heart, (*b*) widespread vasodilatation, and consequently (*c*) fall of blood pressure (Fig. 477), (*d*) diminished adrenaline secretion, (*e*) depression or arrest of respiration, (*f*) diminished tone in skeletal muscle, (*g*) visceral changes, *e.g.* increased gastric tone and movement, and decreased bladder tone (Fig. 478) resulting from the altered level of activity of the autonomic nerve supply of these organs.

(ii) Study of the sinus nerve action potentials shows that as the endo-sinusal pressure rises their frequency increases (Fig. 479). As many more

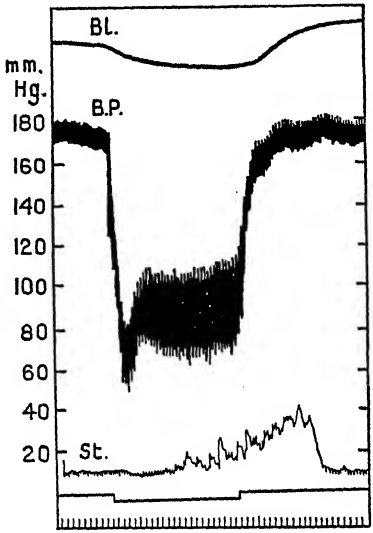


Fig. 478.—Reflex Effects of Raising Carotid Sinus Pressure on Bladder, Stomach, and Blood Pressure. (Schweitzer, *Pflug. Arch.*, 1934, *235*.)

Records from above downwards are urinary bladder pressure (Bl.); blood pressure (B.P.); movements of stomach (St.); signal, time in seconds. At the signal the isolated innervated carotid sinus was distended at a pressure of 230 mm. Hg. Note reflex fall of blood pressure and cardiac slowing (greater excursions of mercury), diminution of bladder tone and increase in gastric tone and movements.

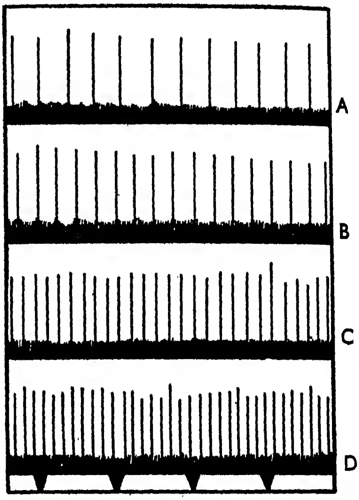


Fig. 479.—Afferent Impulses from a Single End Organ in the Carotid Sinus Stimulated by steady Pressure within the Sinus. (Bronk and Stella, *Amer. J. Physiol.*, 1935, *110*.)

In A the sinus pressure is 40 mm. Hg
,, B ,, ,, ,, 80 ,, ,,
,, C ,, ,, ,, 140 ,, ,,
,, D ,, ,, ,, 200 ,, ,,

Note the increase in the frequency of the discharge as the pressure in the sinus is raised.
Time marker = 0·2 seconds.

impulses thus arrive at the bulbar centres per unit time, they should produce the same effects as increasing the frequency of artificial stimulation of the nerve, *i.e.* a greater reflex depression of the circulation and greater alterations of other affected functions; such indeed is the result observed.

(iii) When the pressure in the sinuses is artificially *lowered* the reverse effects are produced, *i.e.* cardiac acceleration, increased cardiac excitability and development of irregularities due to extra-systoles, vasoconstriction, rise of blood pressure, increase in rate and depth of respiration, increased adrenaline secretion and changes in gastro-intestinal tone and movements. The action potentials diminish in frequency. It should be emphasized that

all the physiological effects described are due to removal of the tonic inhibitory afferent impulses in the sinus nerves.

(iv) *Occlusion* of the common carotid arteries in the intact animal produces similar results (Fig. 480, A). These are not due to the associated cerebral

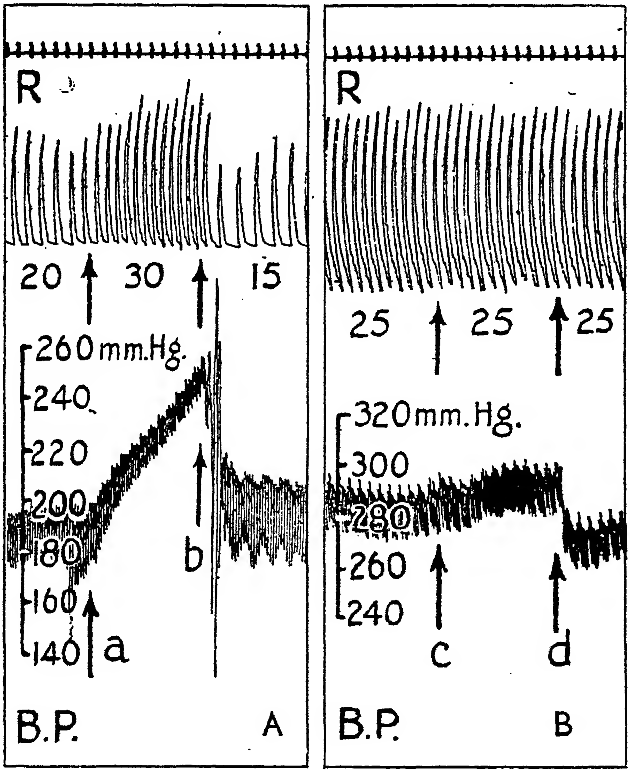


FIG. 480.—Reflex Effects of Occlusion of Common Carotid Arteries on Blood Pressure Heart Rate, and Respiration (Carotid Sinus Reflexes). (Heymans and Bouckaert *J. Physiol.*, 1931.)

Experiment on dog: chloralose. Aortic nerves cut.

Records from above downwards: Time in 3 seconds; R: respiration (inspiration = upstroke); figures below record denote rate of breathing per minute; B.P.: arterial blood pressure.

A. Between *a* and *b* occlude common carotid arteries: fall of pressure in carotid sinuses produces increase in rate and depth of respiration, rise of arterial blood pressure, and increase in heart rate. At *b*, release arteries: sudden distension of carotid sinuses reflexly causes temporary inhibition of breathing, slowing of heart, and fall of blood pressure.

Both carotid sinus nerves cut between A and B. Blood pressure rises from 190 to 280 mm. Hg, and heart accelerates; breathing increases in rate and depth.

B. Between *c* and *d* repeat occlusion of common carotid arteries: no change in respiration or heart rate; slight mechanical increase of blood pressure. At *d* release arteries: slight mechanical fall of blood pressure.

anæmia. If the arteries are occluded following section of the sinus nerves, the changes produced are negligible (Fig. 480, B).[1]

[1] When the blood flow through the carotid body is reduced the *chemoreceptors* are stimulated, reflexly raising blood pressure; this factor contributes to the rise of blood pressure produced by lowered carotid sinus pressure or carotid occlusion (p. 746).

The high degree of sensitivity of the carotid sinus nerve endings to small internal pressure changes is discussed on pp. 273, 312.

(4) *Effects of Section of Sino-Aortic Nerves.*—(i) *Heart Rate.*—Section of the sinus or aortic nerves alone usually results in cardiac acceleration. This proves that normally they exert a tonic afferent inhibitory influence on the resting heart rate ; in terms of the main part of the cardiac centre this means that these afferent nerves reflexly, tonically *stimulate* the cardio-inhibitory centre and increase the discharge of impulses along the vagi to the heart.

All four vaso-sensory nerves may be removed simultaneously in acute experiments, or by means of a two-stage operation in animals that are allowed to survive, so that the chronic effects can be observed. The heart accelerates to the same degree as if the vagi had been divided ; subsequent section of the vagi produces no further acceleration. These experiments

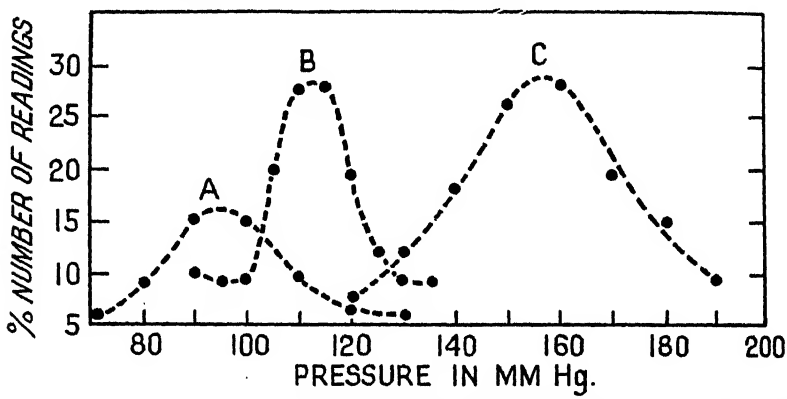


FIG. 481.—Chronic Hypertension in the Rabbit following Division of the Sino-aortic Nerves. (Kremer, Scarff, and Wright, *Brit. J. exp. Path.*, 1933.)

A = Normal range of blood pressure in rabbit (mean 90 mm. Hg).
B = Range of B.P. after division of sino-aortic nerves on *one* side (mean 110 mm. Hg).
C = Range of B.P. after division of sino-aortic nerves on both sides (mean 155 mm. Hg).

prove that *resting vagus tone* is wholly *reflex* in origin and depends on afferent impulses along the sinus and aortic nerves (cf. p. 270).

In these " deafferented " animals marked cardiac irregularity may be present, usually consisting of runs of *extra-systoles* arising from the right or the left ventricle, or from both. In acute experiments fatal *ventricular fibrillation* or pulmonary œdema often develops. The irregularity is abolished by section of the cardiac sympathetics or their paralysis with ergotamine; the cardiac irregularity is thus due to overactivity of the sympathetic innervation of the heart. These results indicate that the sino-aortic nerves normally exert a tonic (afferent) inhibitory effect on the cardiac sympathetic and thus help to maintain the *normal cardiac rhythm.*

(ii) *Blood Pressure.*—If both sino-aortic nerves are cut in an acute experiment the blood pressure rises markedly, owing to increased vasoconstriction resulting from release of the vasomotor centre from tonic inhibitory impulses, and greater adrenaline secretion. If the nerves are removed in two stages (separated by several weeks' interval) the animals may survive for long periods

and develop *persistent hypertension.* Experiments along these lines have been performed successfully in the rabbit, cat, dog, and monkey, and with similar results. In the rabbit, for example, the average normal blood pressure is 90 mm. Hg.; after section of one pair of afferent nerves it is 110 mm. Hg, and after bilateral denervation, 155 mm. Hg (Fig. 481). The blood pressure tends to fluctuate spontaneously far more widely than in intact animals (cf. p. 345).

Rôle of Chemoreceptors (Carotid and Aortic Bodies).[1]—(1) *Carotid Body.*—Using the perfusion technique already described it can be shown that varying the chemical composition of the blood or other fluid flowing through the region of the carotid bifurcation reflexly affects respiration, heart rate, and blood pressure.

(i) If the perfusing fluid contains *excess CO_2*, has a raised *H^+ ion concentration*, or is *deficient in oxygen*, the principal effect is that respiration is reflexly stimulated. The reflex stimulating action of excess CO_2 is well shown in Fig. 482. The frequency of the action potentials in the chemoreceptor fibres of the carotid sinus nerve trunk, simultaneously rises indicating that more nerve impulses are being sent up to the respiratory centre. Commonly there are also reflex effects on the circulation, *e.g.* acceleration of the heart and a rise of blood pressure (cf. p. 746). Conversely, if the perfusing fluid is *deficient in CO_2* or is *too alkaline*, both respiration and circulation are reflexly depressed.

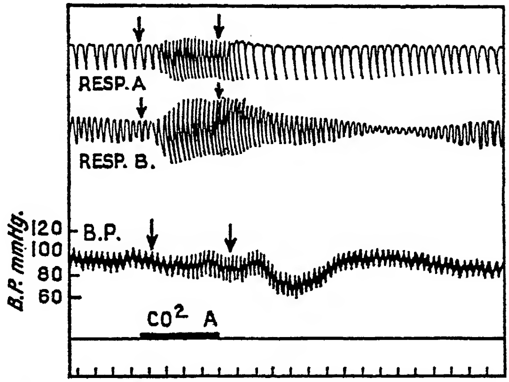


FIG. 482.—Reflex Stimulation of Breathing by CO_2 via Chemoreceptors in Carotid body. (Gayet, Bennati, and Quivy, *Arch. internat. Pharmacodyn.*, 1935, *50.*)

Records from two dogs, prepared according to technique illustrated in Fig. 461, *i.e.* carotid bifurcation of B is perfused with blood from carotid artery of A. Records from above downwards are: respiration in A (donor); respiration in B (perfused from A); blood pressure in B; signal; time in 10 sec.

Between the arrows the donor (A) inhaled a CO_2 rich mixture. Note stimulation of breathing in A and *reflex* stimulation of breathing in B.

(ii) Certain drugs of physiological interest may produce reflex stimulation of breathing. Fig. 483 shows a striking increase of respiration produced reflexly by the action of *acetylcholine* on the chemoreceptors in the carotid body. *Cyanide* (which may act by inducing local anoxia of the nerve endings) and *nicotine* act similarly.

(iii) *Localization of Carotid Chemoreceptors.*—The chemoreceptors are situated in the carotid body. In several species the carotid body receives its blood supply from the occipital artery; if this vessel is tied the carotid sinus and its innervation is left intact, but no blood can reach the carotid body or act on its sensory endings. If the carotid bifurcation is now artificially

[1] Symposium, *Acta physiol. scand.*, 1951, *22*, 4–82. Schweitzer and Wright, *Quart. J. exp. Physiol.*, 1938, *28*, 33.

perfused it responds normally to changes of internal *pressure* (*i.e.* pressoreceptors intact), but no reactions are obtained when the chemical composition of the blood is altered (*i.e.* chemoreceptors not functioning).

(2) *Aortic Chemoreceptors.*—The rôle of the *aortic* chemoreceptors can be demonstrated by more elaborate procedures. An animal is prepared as described in the legend to Fig. 484. If the trunk is asphyxiated (Fig. 484) or is subjected to CO_2 excess or oxygen lack, or if acid is injected intravenously into it, breathing in the head is *reflexly* stimulated; if the trunk is overventilated (to wash out CO_2), or if alkali is injected, the breathing in the head is *reflexly* decreased or even stopped. It must be remembered that in these experiments the composition of the blood supply to the respiratory centre has not been altered at all. The responses described are not abolished by section of the pulmonary fibres of the vagus; the sensory endings are in the *aortic body*.

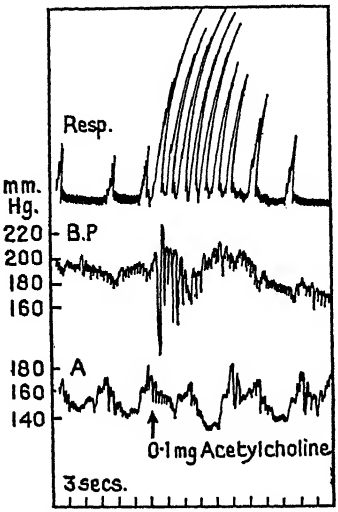


FIG. 483.—Reflex Effects of Acetylcholine on Respiration. (Heymans and Bouckaert, *Ergeb. Physiol.*, 1939, *41.*)

The carotid sinuses of the recipient dog B were perfused by the technique illustrated in Fig. 461 by blood from a donor dog A. Records from above downwards are: respiration of B (recipient); blood pressure of B; blood pressure of A (perfuser); time in 3 seconds. At arrow, inject 0·1 mg. acetylcholine into carotid artery of perfuser. Note intense reflex stimulation of breathing and slowing of the heart in B.

(3) Stimulation of the aortic or carotid chemoreceptors by oxygen lack causes powerful reflex stimulation of the *vasomotor centre* which contributes largely to the rise of blood pressure caused by anoxia (p. 310). The chemoreceptors are also stimulated when the *blood flow through them is reduced.*[1]

In *hæmorrhage*, the blood pressure tends to fall, reducing the tonic inhibitory influence of the sino-aortic nerves on the vasomotor centre. In addition, owing to the decreased cardiac output (and possibly because of constriction of the blood vessels to the carotid and aortic bodies) (Neil) the blood flow to the chemoreceptors is reduced with resulting increase in their activity; the afferent impulses set up reflexly *stimulate* the vasomotor centre. The vasomotor centre is thus: (i) *released* from tonic afferent inhibition; (ii) reflexly *stimulated*; general vasoconstriction results which helps to maintain or restore the blood pressure after hæmorrhage (p. 82). When the carotids are occluded asphyxia of the carotid body results. The chemoreceptor discharge is increased and contributes to the reflex rise of blood pressure which occurs (p. 743).

Afferent Nerves from the Lungs.—The vagus nerve supplies afferent fibres to the larynx, respiratory passages, and the alveoli. These fibres are protective, and also serve to regulate the rhythm of respiration (p. 387). When the lungs are altered by disease (congestion, consolidation, or collapse) afferent impulses may be set up which reflexly stimulate breathing (pp. 457, 458).

Afferent impulses from the lungs themselves do not give rise to conscious

[1] Neil, *Acta physiol. scand.*, 1951, *22*, 54.

sensations. Extensive pulmonary disease which does not involve the parietal pleura may be wholly painless.

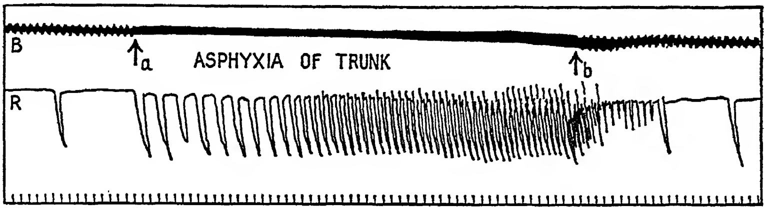


FIG. 484.—Reflex Stimulation of Respiratory Centre from Aortic Chemoreceptors.

B = Blood pressure of trunk; R = Respiratory record of head. The head is isolated from the trunk and only the vagi are left connecting the two regions. The head is kept alive by perfusion with blood from another animal. The trunk is kept alive by artificial respiration. Between "*a*" and "*b*" the artificial respiration was stopped in the trunk of B. Asphyxia of the trunk reflexly stimulates (via the vagi) the respiration in the head. (Heymans and Heymans, *Arch. internat. Pharmacodyn. Therap.*, 1927.)

Cutaneous Hyperalgesia (Tenderness).[1]—The subject is discussed a this point because of the light it throws on the understanding of some of the phenomena of visceral disease. A strong faradic current, of an intensity to produce almost unbearable pain, is applied to the skin (*e.g.* on the surface of the forearm) for 5 minutes; the current injures the skin, as shown by the appearance of a wheal (p. 324). Subsequently the surrounding skin becomes sore or tender (*cutaneous hyperalgesia*). In the affected area, the sense of touch (to stimulation with cotton wool or Frey's hairs) is unaffected or diminished, but a needle prick produces an intense, diffuse, long-lasting pain; the skin is excessively sensitive to slight friction. A certain degree of spontaneous burning, smarting, or itching may be present. This cutaneous hyperalgesia increases in intensity, reaching its maximum in about 20–30 minutes; it also spreads progressively up and down the arm and may ultimately involve an area of some 20 square inches. The tenderness persists for hours, or even for as long as a day or more, and gradually disappears. The degree and extent of the hyperalgesia vary considerably in different individuals. Similar results are obtained if the skin is injured by other procedures, *e.g.* freezing or pinching.

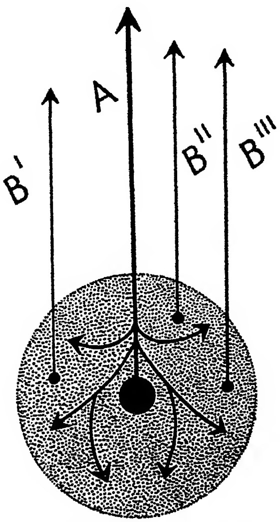


FIG. 485.—Possible Neural Basis of Cutaneous Hyperalgesia.

Black circle: skin area subjected to strong faradic stimulation. Large dotted circle: area of cutaneous hyperalgesia.

Analysis shows that the following mechanism is involved (Fig. 485). When the skin is injured chemical agents are released which stimulate the local nociceptive (pain) endings; impulses consequently pass up along afferent nerves (A) to the central nervous system and give rise to the initial sensation

[1] Lewis, *Clin. Sci.*, 1936, *2*, 373; 1937, *3*, 59.

of localized pain. It is known that the cutaneous afferent nerves branch in the skin to supply blood vessels in the vicinity; it is suggested that in addition other branches are given off which ramify over a considerable area, ending throughout in *skin cells.* It is thought that when impulses along these filaments reach the skin cells they release some stable (and unidentified) chemical product which modifies the behaviour of the local cutaneous nerve endings, *e.g.* those connected with B′, B″, B‴, so giving rise to the abnormal state of skin sensibility already described. We seem to be dealing with the *neural transmission of a chemical disturbance* (the converse of the humoral transmission of a nervous disturbance (p. 508)); *i.e.* a chemical change set up in a discrete area of skin by an injury sets up nervous impulses which in their turn produce a long lasting chemical disturbance in a related wide skin area. In this latter zone the local pain nerves "register" a skin condition which is *not directly produced by local external stimulation.* The mechanism described resembles the axon reflex which is responsible for the "flare" in

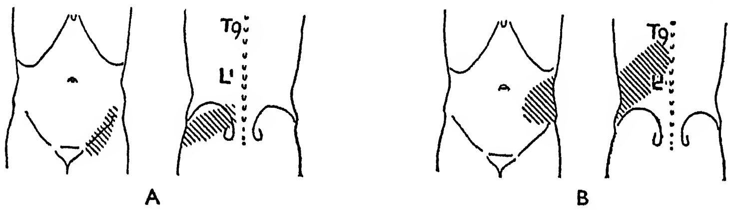


FIG. 486.—Areas of Cutaneous Tenderness Following Injection of 6% Saline into (A) left first Lumbar, and (B) left ninth Thoracic Interspinous Ligament. (Lewis and Kellgren, *Clin. Sci.*, 1939, *4*, 48.)

injured skin because the hyperalgesia too depends on the branching of nerve fibres. The wider significance of these results is considered on p. 759.

Referred Pain, Tenderness, and Rigidity evoked from Deep Somatic Structures.[1]—If hypertonic saline (6%) is injected into an *interspinous ligament, pain* is set up, which is *referred to the peripheral distribution of the corresponding segmental nerve.* The pain is described as a continuous ache, lasting for minutes, which is felt deeply in the referred zone; it is accompanied by tenderness and muscular rigidity.

Three striking observations may be quoted.

(1) The *first lumbar* interspinous ligament is injected. The pain is felt in the loin, inguinal, and scrotal region; it may last for minutes and be accompanied by retraction and tenderness of the testis and localized *rigidity* and *deep* tenderness in the lowest part of the abdominal wall. The distribution of the *cutaneous* tenderness is shown in Fig. 486; it does not correspond exactly in area with the deep tenderness. The symptoms and signs produced resemble very closely those of *renal colic* which are set up when a stone is passing down the ureter (p. 763, footnote 1).

(2) The *eighth cervical* interspinous ligament is injected. The distribution of the resulting pain is shown in Fig. 487. The pain is accompanied by a sense of constriction in the upper part of the chest on the affected side. If the ligament on the left side is injected in patients who suffer from angina of effort (p. 752), the pain experienced is affirmed by them to be identical

[1] Kellgren, *Clin. Sci.*, 1939, *4*, 36. Lewis and Kellgren, *ibid.* 47.

in character with that experienced during an attack, though it differs somewhat in distribution. The "injection pain" is felt mainly in the back, while the anginal pain is felt mainly (though by no means invariably) in the front of the chest.

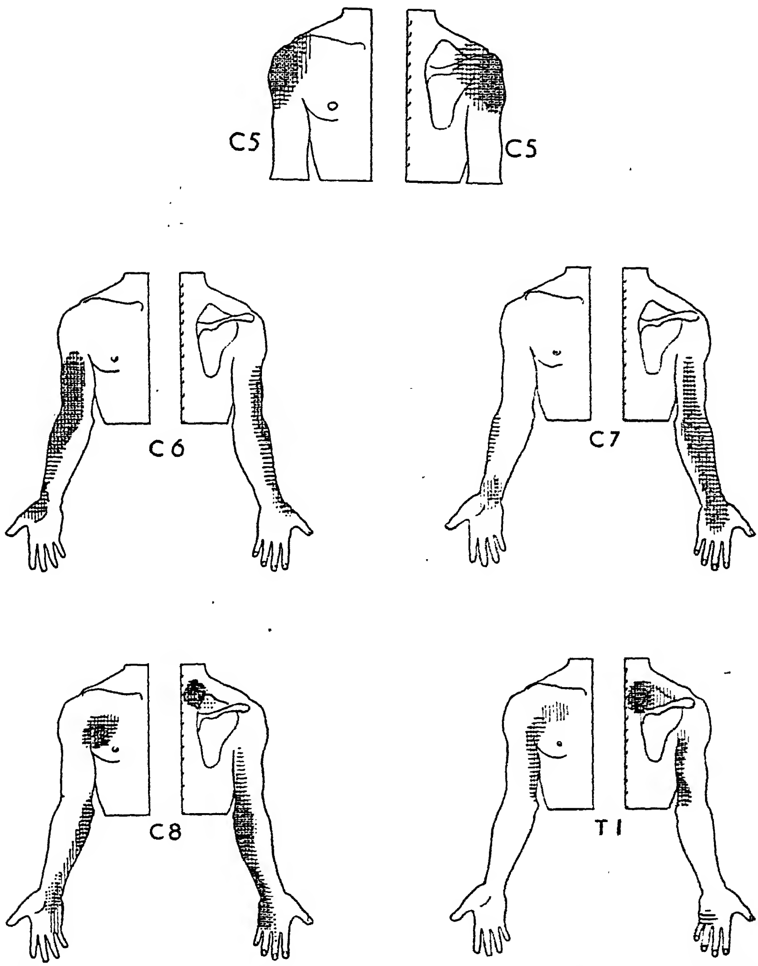


FIG. 487.—Distribution of Pain following Injection of 6% Saline into Interspinous Ligaments of C5, C6, C7, C8, and T1. (Kellgren, *Clin. Sci.*, 1939, *4*, 37.)

Each chart shows results in three subjects (vertical hatching, horizontal hatching, and stippling).

(3) Injection of the *third thoracic* ligament in the middle line gives rise to bilateral pain over the sternum and a sense of constriction.

Fig. 487 also shows the areas of referred pain resulting from stimulation of the interspinous ligaments of C5, 6, 7, 8, and T1.

According to Lewis, stimulation by means of hypertonic saline of *superficial*

somatic structures, such as deep fascia, subcutaneous ligaments, tendon sheaths (*e.g.* tendo Achilles), or subcutaneous periosteum (*e.g.* over the tibia) gives rise to pain which is fairly accurately *localized* in the region affected. Stimulation of *deep* somatic structures gives rise to referred pain.

Fig. 488, which illustrates these points well, shows the effects of stimulating the various structures in the chest wall at the level of the sixth intercostal space. Stimulation of the deep fascia and periosteum gives rise to localized pain. Stimulation of the deeper-lying muscles, on the other hand, gives rise to referred pain which is felt in the *skin distribution of the segmental innervation* of the muscles concerned. Thus stimulation of the trapezius muscle gives rise to referred pain in the skin distribution of C3 and 4 and the accessory nerve.

Pain of Cardiac Ischæmia (Coronary Insufficiency).—Angina Pectoris.—Severe pain of cardiac origin is usually (and probably invariably)

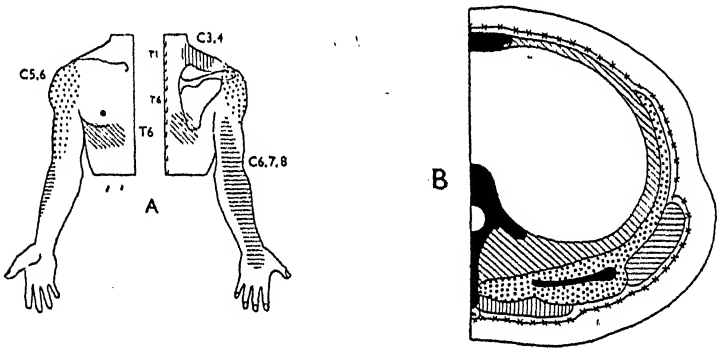


FIG. 488.—Referred Somatic Pain. (Kellgren *Clin. Sci.*, 1939, *4.*)

A. Distribution of pain arising from various deep structures of the chest wall at the level of the sixth intercostal space. Pain produced by local injection of 6% saline.
B. Tissues which give rise to corresponding areas of pain.
Oblique hatching: intercostal space and erector spinæ (T6).
Stippling: muscles attached to scapula (C5, 6).
Horizontal hatching: latissimus dorsi (C6, 7, 8).
Vertical hatching: trapezius (accessory nerve and C3, 4).
Crosses: tissues giving rise to *local* pain (not shown in A).

the result of an inadequate blood flow (*ischæmia*) to the myocardium. Thus, very violent and prolonged pain is set up when a branch of one of the coronary arteries is occluded as a result of thrombosis (*coronary thrombosis*, cf. p. 239),

(1) The *pain* (so-called *anginal pain*) may persist for hours, or even days, is agonizing in character, but is *not felt in the heart.* It is characteristically referred to the *sternum*, the outer part of the left side of the chest (Th1, 2, 3), down the inside of the arm (Th1), up the neck to the angle of the jaws (C3, 4), and to both shoulders and back: the pain may be referred to the epigastrium (simulating that of an abdominal catastrophe) and give rise to vomiting. There is thus in this instance *no local visceral pain* (*i.e.* no pain felt *in the viscus*) but only a referred pain. A striking feature of coronary thrombosis is the very wide area involved by the referred pain, wider than corresponds to the sensory segmental supply of the heart. In other words, a form of *sensory irradiation* takes place to adjacent segments.

(2) *Tenderness* is present over these regions too.

(3) There is a sense of constriction, as though the chest were gripped in a vice, or as if the sternum would break. Mackenzie believed that owing to a *viscero-motor reflex* there is reflex spasm of the intercostal muscles, which prevents respiratory movements of the chest.

The effective stimulus to the nerve endings in the heart in this condition is not pressure, but (as shown below) the *chemical products of ischæmia.*

MECHANISM OF ISCHÆMIC PAIN.[1]—Lewis studied the causation of pain which develops in a limb when the circulation is entirely occluded. Exercise is carried out by performing a gripping movement with the hand at the rate of once per second. Pain sets in after 30 seconds and is intolerable in 70 seconds. Though the pain is diffusely felt, it is most marked in the *muscles.* Various possible causes of the pain may be considered.

(i) It is not due to *vascular spasm,* because after occlusion the blood vessels of a limb lose their tone.

(ii) The pain is not due to *muscular tension,* because it is continuous and not accentuated during contraction.

(iii) It is the *result of activity* because it is related to the amount of exercise which is performed.

(iv) The following observations show the pain is not due simply to local anoxia : Exercise the arm till pain develops and note the time taken ; allow a suitable interval for recovery to occur, and then repeat the exercise, but for a few seconds less ; this time no pain develops. Now maintain the occlusion for another 5 minutes. The oxygen content of the limb must decrease markedly during this period but no pain develops.

Lewis suggests that muscular activity releases a pain-producing factor (P) which passes out into the tissue spaces and is normally removed by the blood stream. If exercise is carried out with the circulation occluded this substance accumulates, and when it reaches a certain concentration, pain develops. In support of this view, Lewis finds that if exercise is performed with occluded circulation till pain appears, and the exercise is stopped but the occlusion maintained, the pain persists unchanged and *does not get worse* ; if the circulation is released, the pain disappears within a few (2–4) seconds.

INTERMITTENT CLAUDICATION.—The above experiments explain the recurrent pain produced in the legs during exertion in patients with narrowed limb blood vessels (*e.g.* in *Buerger's disease*) [2] (*intermittent claudication*). In this disease, the blood supply to the muscles is adequate for their needs during periods of rest ; but during activity the blood flow cannot be increased sufficiently to cope with the additional requirements (*i.e.* there is *relative* ischæmia). The P factor consequently accumulates, giving rise to pain which increases in intensity until the patient is compelled to stop. During the period of rest the P factor is washed away, the pain disappears, and the patient can resume walking, only to be stopped again by recurrence of the pain.

Other Phenomena in Ischæmic Limb.[3]—If a blood pressure cuff wrapped round a limb is inflated to occlude the underlying artery for 15 minutes, the nerve trunks *under the cuff* begin to lose their conductivity. The fibres first affected are those which have travelled the longest distance from their nerve

[1] Lewis, *Arch. int. Med.*, 1932, *49*, 713. Levy, *Diseases of Coronary Arteries and Cardiac Pain*, N.Y., 1936.

[2] Also called *thromboangiitis obliterans* (cf. p. 361).

[3] Lewis, Pickering, and Rothschild, *Heart*, 1929–31, *15*, 359 ; 1931–33, *16*, 1.

endings; the fibres which have travelled a shorter distance are affected later (cf. p. 1005). After initial tingling a defective sense of touch develops in the most distal part of the limb; it gradually spreads up proximally and is followed by a similar centripetal loss of other forms of sensation, pain being lost last. Loss of motor function shows a like march and is almost simultaneous with loss of touch sense.[1]

Pain of Coronary Occlusion.—The pain of coronary occlusion, like that in an ischæmic limb, is presumably produced by ischæmia of the affected area of heart muscle. The pain persists as long as the sensory nerve endings in the ischæmic patch of heart muscle remain alive; when they die the pain disappears unless spasm of other coronary vessels takes place (*infra*).

Angina of Effort.—Attacks of anginal pain may be brought on by *effort*, without evidence of coronary occlusion. Little is known about the changes in the heart rate or blood pressure during the attacks, but they always occur under circumstances that would be expected to throw an extra burden on the heart and so cause *relative* ischæmia. If there is spasm or anatomical narrowing of the coronary vessels the necessary degree of local vasodilatation may not occur in conditions of stress; the blood supply of the heart is thus insufficient for its increased needs, the P factor is not got rid of, it accumulates and gives rise to pain. Severe anginal pain may, however, occur at night or at other times when the subject is at complete rest. If there is no evidence of occlusion the pain must be attributed to coronary spasm reducing the cardiac blood flow.

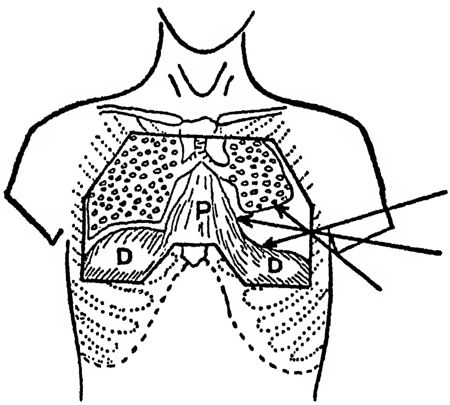


Fig. 489.—Diagram illustrating the Method of Testing Pleural Sensibility. (Capps, *Experimental and Clinical Study of Pain*, New York, Macmillan & Co., 1932.)

The left lung is displaced by effusion and the diaphragm is pushed down. The arrows represent the wire passing through the hollow trocar and touching respectively the pulmonary, pericardial (P), and diaphragmatic pleura (D).

Palpitation.—When consciousness of the heart's action is present, it is called palpitation. It is associated with vigorous exercise or with the strong contraction which often follows on a premature beat. The sensation is set up by the impact of the heart against the chest wall.

Pain of Arterial Embolism and Thrombosis.[2]—Pain from this cause only develops in muscular organs, *i.e.* the *intestine*, *heart*, and *skeletal muscle.* Interference with the blood supply to the lungs, liver, brain, or spleen does not give rise to pain. Variable results are obtained in the case of the kidney; possibly pain only develops if the muscular renal pelvis is involved.

Pleural Sensibility.—The experimental study of pleural sensibility in man yields results which illustrate important principles which apply to

[1] For results of asphyxia and pressure on sensory nerves, see Lewis and Pochin, *Clin. Sci.*, 1938, *3*, 141.

[2] Lewis, *Clin. Sci.*, 1935–36, *2*, 238.

other *serous membranes*, certain *viscera* and also certain *deep somatic structures*.

This problem was studied by Capps in patients with pleural effusion in the following way (Fig. 489). A hollow trocar is pushed through the chest wall into the pleural cavity; before much fluid escapes, a long stiff silver wire is passed through the cannula and brought in contact with various aspects of the pleural membrane. When the pleura was covered with thick exudate, no response was obtained; but when the pleura was in a more normal condition, the following results were elicited:

(1) Visceral Pleura.—The *visceral* pleura (lining the lung) is entirely devoid of pain sense.

(2) Parietal Pleura.—Pressure on the *parietal* pleura (lining the chest wall) gives rise to sharp pain which is located with considerable accuracy near the site of stimulation. The parietal pleura can be regarded as part of the chest wall and like it is supplied with sensory fibres from the intercostal nerves.[1]

(3) Diaphragmatic Pleura (see Fig. 490). (i) The *central* portion of the diaphragm is supplied

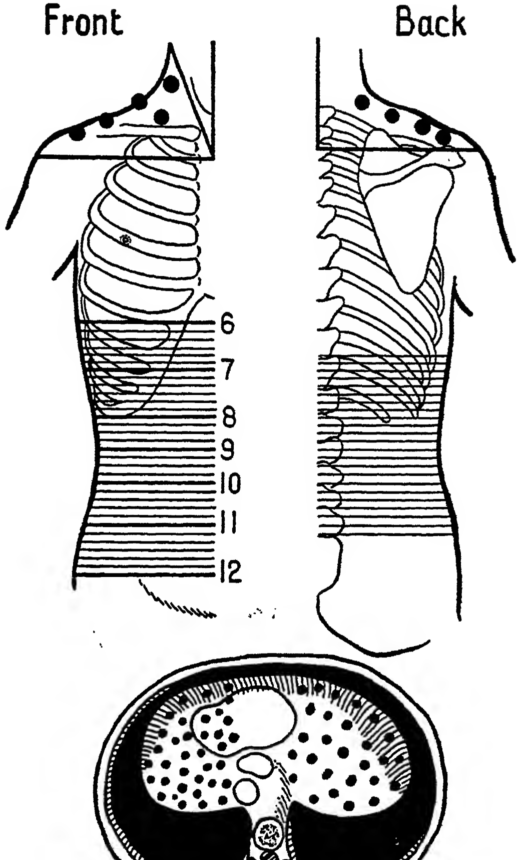


Fig. 490.—Localization of Pain from Diaphragmatic Pleura (or Peritoneum). (Capps, *Study of Pain*, 1932.)

The lower diagram represents the thoracic aspect of the diaphragm. The area marked with black dots is supplied by the phrenic, that shown as a black sheet is supplied by the lower intercostal nerves. Stimulation of the pleura over the central region of the diaphragm gives rise to referred pain in the neck (indicated by black dots in upper figure) in the distribution chiefly of the third and fourth cervical dorsal nerve roots. Stimulation of the peripheral diaphragmatic pleura produces pain referred to the body wall (front or back, of the same side or both sides) in the distribution of the sixth to twelfth thoracic nerves (horizontal lines in upper figure). ***Exactly similar results are obtained on stimulation of the peritoneal surface of the central and peripheral parts of the diaphragm.***

[1] The parietal pleura can sometimes be incised without producing pain when rib resections are carried out under local anæsthesia; presumably the pathological condition present in the pleura abolishes its normal sensitivity. Healthy subjects usually complain of pain when an exploring needle comes in contact with the pleura.

with afferent (as well as efferent) fibres by the phrenic nerve. Experimental irritation of this region sets up a sharp, cutting pain in a region anatomically quite remote from the diaphragm, namely in the *neck*—usually in a definite spot—in the skin distribution of the third and fourth cervical dorsal nerve roots. This is a clear-cut instance of *referred pain* set up from a serous membrane. The phrenic afferents enter the dorsal nerve roots of the third and fourth cervical nerves mainly (and also C5). The patient is unable to localize accurately afferent impulses that have come from the diaphragmatic pleura, and falsely projects—refers—the sensation to *a region of the body wall which is supplied by the same dorsal nerve roots as supply the diaphragmatic pleura.* Referred pain is a special form of *false localization.* The pain in the neck is

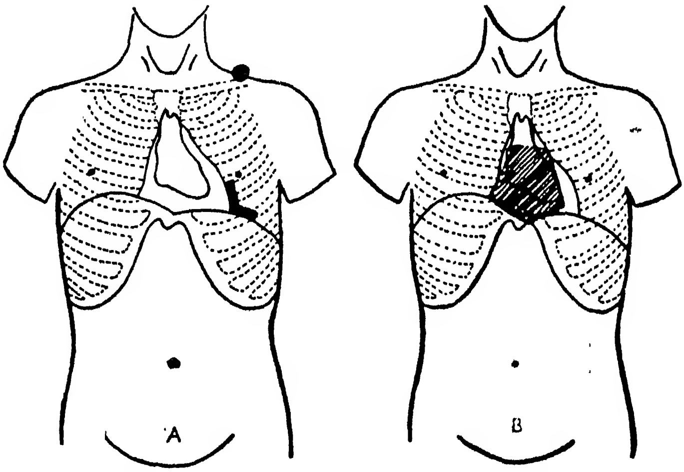


FIG. 491.—Pericardial Pain. (Capps, *Experimental and Clinical Study of Pain*, New York, Macmillan & Co., 1932.)

A. Heavy line shows portion of fibrous pericardium and central diaphragm, irritation of which gives rise to neck pain (black circle).
B. Case of typical dry pericarditis with to-and-fro friction sounds over the whole area of the heart (shaded area). Post-mortem examination showed universal fibrinous exudate over the visceral and parietal layer of the *serous* pericardium. *No pain* was felt in the region of the heart at any time.

accompanied by local *tenderness* and sometimes by reflex *rigidity* of the neck muscles.

(ii) The *peripheral* part of the diaphragm is supplied with afferent (as well as efferent) fibres from the lower six intercostal nerves (Th7–12). Experimental irritation of the pleura covering this region gives rise to pain which seems to come from (*i.e.* is referred to) a varying number of segments of the body wall supplied by the lower thoracic nerves—the lower thorax, the lumbar region, or the abdomen. This is another clear-cut instance of referred pain. The stronger the degree of pleural stimulation the more widespread is the referred pain; it may sometimes extend to the *other side* of the body too (sensory irradiation). The pain is associated with local *tenderness* and reflex motor effects in the form of *muscular rigidity.*

PAIN OF PLEURISY.—When the parietal pleura is involved in inflammatory processes, the pain is felt in the chest wall over the region involved; with

diaphragmatic pleurisy, the pain is referred to the neck, or the lower part of the chest, the back, or abdomen. When the pain is referred to the abdomen, the clinical findings may lead to a wrong diagnosis of an intra-abdominal catastrophe.

Pericardial Sensibility.—The studies of Capps have led to the following conclusions.

(1) The lower part of the *fibrous* pericardium (at the level of the fifth and sixth interspaces) is supplied by afferents from the phrenic nerve; puncture of the pericardium in this region gives rise to referred pain in the *neck* (Fig. 491, A).

(2) On the other hand, experimental stimulation of the *serous* pericardium (both visceral and parietal layers) gives rise to no pain. A *dry* inflammation of the serous pericardium may produce a sense of precordial tightness, but again, as a rule, no pain (Fig., 491, B). A large *effusion* greatly stretching the pericardial sac may give rise to no more than a dull ache over the heart.

(3) Severe pain is felt (i) when the inflammatory process involves the *mediastinal tissues* (the pain is felt over the heart); or (ii) the *diaphragmatic pleura* (the pain is then referred to the neck or abdomen); or (iii) if the *coronary circulation* is inadequate: the pain is then due to cardiac ischæmia (p. 750).

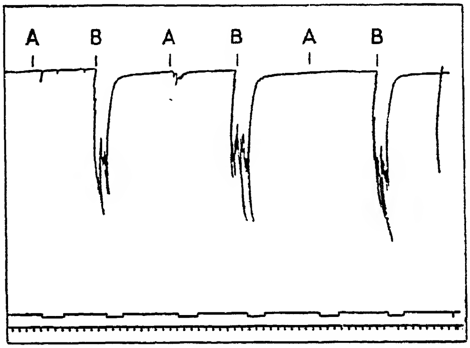


Fig. 492.—Reflex Contraction of Abdominal Wall from Stimulation of Diaphragm. (Lewis and Kellgren, *Clin. Sci.*, 1939, *4*, 56.)

Records from above downwards: Contraction of right upper rectus; signal line; time in secs. Stimulate the under surface of diaphragm (right side) with a smooth probe during the period shown by the signals. A. Stimulate dome (no reflex response); B. stimulate lateral part of diaphragm (reflex contraction).

Peritoneal Sensibility.—The parietal peritoneum and adjacent connective tissue are supplied with *somatic* afferent fibres from the lower thoracic nerves.

(1) Diaphragmatic Peritoneum.—Using the same technique as for the pleura, Capps has shown that strong pressure, or the application of a rough point to the peritoneal surface of the diaphragm gives rise to referred pain which is felt in the neck when the central parts of the diaphragm are stimulated, and in the chest or abdomen from stimulation of the peripheral parts (Fig. 490). The results are identical with those obtained from the pleural surface of the diaphragm (p. 753) and can be accounted for in the same way. (It should be noted that *light contact or stroking* elicited no reaction.) Similar pain is experienced in inflammatory conditions of this region of the peritoneum. The pain is intensified by forcible movements of the diaphragm such as those produced by deep breathing or coughing. It is associated with local tenderness (hyperalgesia) and rigidity of the neighbouring muscles. Animal experiments show that stimulation of the *peripheral* region of the diaphragm produces *reflex contraction of the abdominal wall* (Fig. 492). The muscular rigidity in clinical cases of irritation of corresponding parts of the diaphragm

is doubtless also reflexly produced.[1] Injury to the spleen associated with irritation of the *central* diaphragm may produce referred pain and hyperalgesia over the shoulder tip (Fig. 493).

(2) Parietal Peritoneum.—Stimulation of the *parietal peritoneum* gives rise to pain which is fairly accurately localized in the superjacent abdominal wall (Fig. 494). Parietal peritoneal pain is also associated with local tenderness and reflex muscular rigidity.

Referred Pain, Tenderness, and Rigidity.—As has already been explained, stimulation of the diaphragmatic pleura or peritoneum, or of the

Fig. 493.—Referred Shoulder Pain from Diaphragm. (Morley, *Abdominal Pain*, E. & S. Livingstone.)

Case of traumatic laceration of diaphragmatic surface of spleen (involving peripheral diaphragmatic peritoneum) showing (as black patch over shoulder) the area of referred pain and hyperalgesia in the skin distribution of the third and fourth cervical nerves.

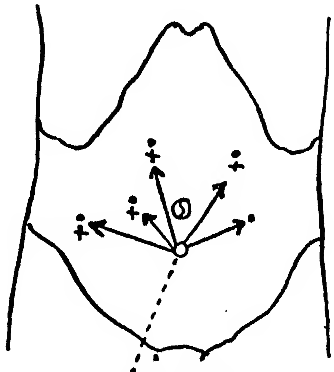


Fig. 494.—Localization of Pain produced by Stimulation of Parietal Peritoneum. (Capps, *Experimental and Clinical Study of Pain*, New York, Macmillan & Co., 1932.)

Anterior abdominal wall: the circle shows where the cannula was inserted through which the wire was passed to stimulate the parietal peritoneum. Black dots=actual point stimulated; + =localization of pain by patient.

interspinous ligaments or other deep somatic structures, or the development of cardiac ischæmia produces referred pain, tenderness (hyperalgesia), and muscular rigidity. The way in which these symptoms and signs are produced is still imperfectly understood and requires discussion.

Referred Pain.—The degree of accuracy of localization of a *tactile* stimulus applied to the skin depends on the local wealth of cutaneous innervation, the extent of the cortical area receiving impulses from the region, and the degree of point-to-point (" topographical ") projection of the receptor surface on to discrete areas of the sensory cortex (p. 570 and Fig. 495). If a sensory surface is sparsely innervated or if many receptors from an area

[1] Stimulation of the back muscles or of the interspinous ligaments (p. 748), produces an identical reflex contraction of the abdominal wall.

converge on to a single cortical point, localization is correspondingly impaired (Fig. 496). As cutaneous pain can also be sharply localized, it is assumed that topographical projection occurs, though the relevant anatomical details are

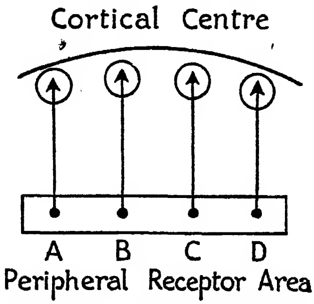


FIG. 495.—Anatomical Basis of Accurate Localization.

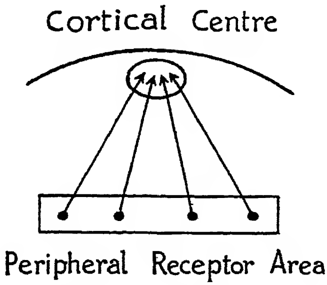


FIG. 496.—Anatomical Basis of Vague Localization.

obscure (p. 571). As the diaphragm, deep somatic structures, and the heart have a relatively sparse sensory innervation, we would expect stimuli applied to these regions to be *vaguely* localized ; what requires explanation is the fact

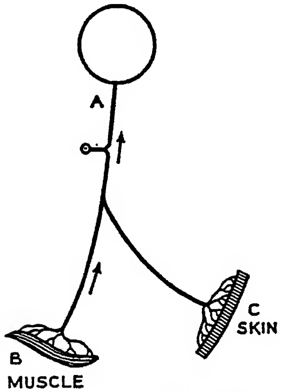


FIG. 497.—Branching of Dorsal Root Axon to supply Skin and Muscle. (Sinclair, Weddell, and Feindel, *Brain*, 1948, *71*.)

Impulses arising from the muscle (B) or the skin (C) would give rise to an identical sensation which would be referred to the " dominant zone " which in this case would probably be the skin.

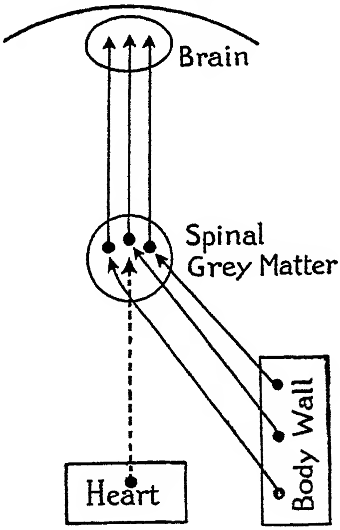


FIG. 498.—Possible Neural Basis of Referred Pain from Heart.

that the sensation is *falsely* " projected " or " referred " to a region remote from the region stimulated. All theories on the subject agree that the false " reference " is due to a *central misinterpretation,* though the suggested detailed explanations vary.

Branching of Dorsal Root Axons.[1]—Referred pain has been attributed to the branching of the peripheral axons of the dorsal roots which convey the impulses set up by nocuous stimuli. It is known that these axons branch extensively in the *skin* mediating the vascular axon reflex which is responsible for the flare (p. 323); a similar ramification is responsible for the cutaneous hyperalgesia which develops round an injured area of skin (p. 747). In addition a few dorsal root axons branch more proximally: (*a*) in the mixed spinal nerve, one branch passing into the posterior ramus and one into the anterior ramus; (*b*) in the anterior ramus, to supply skin and muscle (Fig. 497), skin and a viscus, or skin, muscle, and a viscus (Fig. 500). It is suggested that these branching axons are responsible for referred pain and tenderness.

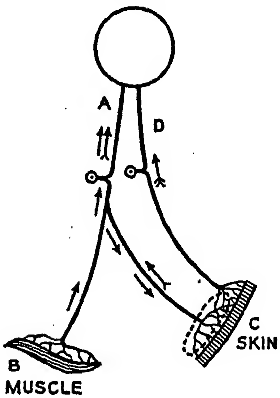


Fig. 499.—Cutaneous Tenderness (Hyperalgesia) set up from an Affected Muscle. (Sinclair, Weddell, and Feindel, *Brain*, 1948, *71*.)

The axon (A) from the muscle (B) gives branches to the skin (C) producing a lasting chemical change there which enhances the sensitivity of the local nerve endings. Gentle stimulation of the skin gives rise to an unpleasant sensation (tenderness) mediated by axons A and D.

Consider the effects of stimulating an interspinous ligament. There is anatomical evidence that the injected hypertonic saline acts on axons in the *posterior* ramus; some of these are processes of dorsal root axons which also contribute a branch to the *anterior* ramus supplying skin and muscle in the anterior part of the dorsal root distribution. As a result, impulses arising from certain points in the posterior or the anterior part of the dorsal root distribution pass into the *same axon*; consequently stimulation of either the anterior or posterior points, or fibres supplying them, must give rise inevitably to the same sensation with respect to quality and localization. We must assume in addition the rule of the "*dominant receptor zone.*" Suppose receptor zone R_1 and R_2 both terminate in the same cortical region C; zone R_1 is frequently stimulated and the mind has consequently "learnt" to project the sensation to the region actually stimulated; zone R_2, on the other hand, is rarely or never stimulated. Then whether zone R_1 or R_2 is stimulated the mind always projects the sensation to zone R_1; *i.e.* impulses from zone R_1 are correctly localized; but impulses from zone R_2 are falsely projected (or referred) to zone R_1. Thus impulses from the interspinous ligaments are falsely referred to the distribution of the *anterior* ramus of the dorsal root employed.

Similar reasoning may be applied to the case of the central diaphragmatic pleura and peritoneum. Some of the afferents from this region are branches of dorsal root axons of C3, 4, and 5, which also supply the skin and muscles of the shoulder region. The latter is the dominant receptor zone; consequently impulses from the central diaphragm are falsely referred to the shoulder region. (Cf. also legends to Figs. 497, 500).

Other Explanations of Central Misinterpretation.—Instead of

[1] Sinclair, Weddell, and Feindel, *Brain*, 1948, *71*, 184.

emphasizing the rôle of branching dorsal root axons it may be supposed that afferents from *e.g.* the upper thoracic body wall and the heart pass along distinct dorsal root axons which end, however, in a common group of cells in the spinal grey matter; the second neurone is common to the receptors zones and ends ultimately in a cortical area which consequently serves both zones (Fig. 498). It is further assumed (as above) that impulses arising from the rarely stimulated zone will be referred to the distribution of the dominant zone.

RIGIDITY.—This is presumably due to reflex contraction of muscles set up by the afferent impulses from the affected area.

TENDERNESS (HYPERALGESIA).—The mechanism may be similar to that of cutaneous hyperalgesia produced by painful skin stimuli. Fig. 499 illustrates how skin tenderness may be set up from an affected muscle. Impulses from the muscle (B) pass along one branch of an axon (A) and then along a second branch to the skin (C) setting up here a long enduring chemical change which enhances the sensitivity of the local nerve endings. When skin (C) is stimulated in a manner which normally would arouse only trivial discomfort (*e.g.* if it is lightly rubbed or pricked) the affected nerve endings set up an unusual pattern of impulses which pass up in the axons (A) and (D) and are centrally interpreted as tenderness. Fig. 500 illustrates how, owing to multiple branching of an axon, stimulation of a viscus may give rise to both cutaneous and muscular tenderness.

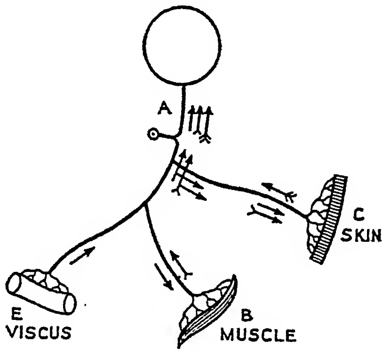


FIG. 500.—Cutaneous and Muscular Tenderness (Hyperalgesia) set up from an Affected Viscus. (Sinclair, Weddell, and Feindel, *Brain*, 1948, *71*.)

The axon (A) receives branches from skin (C) muscle (B), and viscus (E). Impulses reaching the brain from the affected viscus will be referred to the dominant zone which is probably skin (C). Impulses from the viscus (E) pass along the branches of the axon to (B) and (C), producing a lasting chemical change which enhances the sensitivity of the local nerve endings. Gentle stimulation of skin (C) and muscle (B) consequently gives rise to tenderness.

Hyperalgesia has been accounted for by some workers in terms of the "subliminal fringe" (p. 540). Afferent impulses coming from the body wall may stimulate cells in the central grey matter *subliminally* and so not give rise to any conscious sensation. Impulses from an appropriate viscus (set up by disease or experimental stimulation) may be supposed to make connections with the same central pool of neurones. The excitatory state set up from *both* sources may serve to activate the central cells liminally and cause them to discharge, to give rise to an abnormal or exaggerated conscious sensation. This concept is illustrated diagrammatically in Fig. 501.

Abdominal Pain.—Abdominal pain (as already explained) may be produced by stimulation of the parietal or diaphragmatic peritoneum (p. 754) or other deep somatic structures, such as interspinous ligaments or deep muscles (p. 748). We now have to consider pain set up by involvement of the abdominal viscera.

Solid Abdominal Viscera.—(1) LIVER.—Diseases of the *liver* which do not involve the parietal peritoneum may be wholly painless; thus the liver may be extensively involved by cancer without causing pain. In congestive heart failure, however, the liver, which is enlarged and engorged with blood, is usually painful and tender to palpation. It is not easy to explain this difference. It is suggested that when the *capsule* of the liver is stretched the nerve endings there set up the painful impulses; but, on the other hand, stretching of the capsule of the liver enlarged by, say, hydatid disease does not produce pain.

Inflammatory conditions of the liver which reach the surface and involve the parietal peritoneum, stimulate sensory nerves there and, of course, give

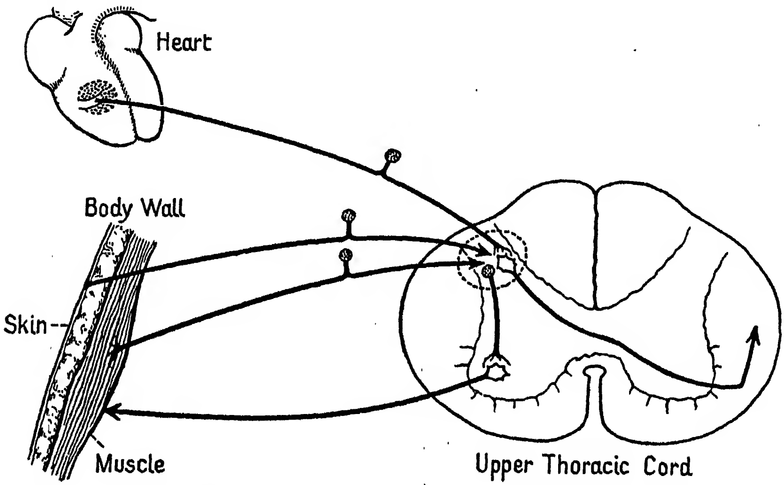


FIG. 501.—Mackenzie's Conception of Neural Basis of Tenderness and Rigidity.

The dotted region in the heart represents a region of myocardium which has become modified as a result of localized coronary occlusion. The region in the dorsal horn of grey matter surrounded by a dotted circle is the region of "subliminal stimulation." Afferent impulses from the heart stimulate these dorsal horn cells, with the result that afferent impulses from the skin and body wall which previously stimulated these cells only subliminally now stimulate them liminally; the afferent impulses are transmitted through these synapses to reach the brain and give rise to a sensation of hyperalgesia. Reflex contraction of muscle can take place. The spinothalamic tract which crosses to the opposite side of the spinal cord conveys the afferent impulses to "consciousness."

rise to pain (p. 756). In animals, experimental stimulation of the liver (or of the kidney or spleen, *v. infra*), *e.g.* by pinching, does not produce reflex muscular contraction.

(2) KIDNEY.—The *kidney* appears to be sensitive to external pressure; when a movable kidney is palpated bimanually, an aching pain is felt locally. It may be that somatic perirenal afferents are here involved. Most diseases of the kidney, *e.g.* nephritis, are quite painless. Stone in the kidney may give rise to a local aching pain, but this may be due to involvement of the perirenal tissues.

The symptoms of *renal colic* are mentioned on p. 763, footnote 1; the afferent impulses probably arise in the muscle coat of the ureter.

(3) PANCREAS.—In the decapitate cat, stimulation of the pancreas (*e.g.* by

pinching) where it lies in the mesentery produces reflex muscular contraction of the abdominal wall; the afferent pathway concerned is in the splanchnic nerves. The afferent endings stimulated may perhaps be those in the mesentery and not in the pancreatic tissue itself. Certain forms of pancreatic disease, *e.g. acute pancreatitis*, are associated with severe abdominal pain and rigidity; others, *e.g.* growth, may be painless during almost all their course.

(4) SPLEEN.—Ischæmia (p. 752) and most other affections of the spleen are usually painless. Pain occurs in rupture of the organ because of involvement of related structures (*e.g.* diaphragmatic (Fig. 493) or parietal peritoneum).

Hollow Abdominal Viscera.[1]—The afferents from these viscera are both non-medullated and finely medullated fibres; they are few in number: thus all the afferent fibres from the viscera of a cat equal in number those in one dorsal nerve root. Because of their sparse innervation, the hollow viscera normally have a low sensitivity; to arouse a sensation, an adequate portion of a viscus must be stimulated with considerable vigour.

There has been much controversy about the sensibility of the hollow viscera, because the clinical study of pain is beset with difficulties and pitfalls. The patient's fear and anxiety may "create" a sensation of pain or an exaggerated reaction to slight pain. The method of questioning, either the tone of the voice or the words used, may influence the patient both as to the presence and absence of pain, its degree, and its situation. A leading question may be indispensable to elicit information but may call forth a misleading reply. In surgical studies under local anæsthesia, the patient's fears may cause him to reply to questions in a manner which he hopes may safeguard him against painful procedures. A local anæsthetic may produce general effects: thus cocaine induces a sense of well-being (euphoria) and damps down unpleasant reactions. The administration of morphine tends to suppress the slighter pains. It seems certain, however, that *adequate* stimulation of viscera may produce pain independently of involvement of the mesentery or the parietal peritoneum.

There is a great difference between the sensitivity of a normal and an *inflamed* viscus, just as there is between normal and inflamed skin. Thus squeezing with finger or forceps or stitching, cutting, or clamping the healthy appendix or cæcum arouses no pain (such observations gave rise to the erroneous conclusion that nocuous stimuli which produce skin pain do not produce pain from abdominal viscera). But when the appendix is inflamed, squeezing it along a considerable part of its length arouses pain. Similarly pinching the normal gastric mucosa or rubbing it with the blunt end of a glass rod is painless; the same stimuli applied to the inflamed and œdematous mucosa arouse pain. The pain fibres in hollow viscera like those in skin or muscle respond to a wide range of stimuli having in common the property of nociception, *i.e.* of being injurious. In discussing cutaneous and muscular hyperalgesia it was suggested that the release of chemical products by axon reflexes modified the sensitivity of local nerve endings thus producing tenderness (Fig. 485). In the viscera, too, states of inflammation by altering the local blood supply or producing œdema or releasing abnormal chemical products may likewise "facilitate" pain responses.

Effects of Raised Tension.—All the hollow viscera are sensitive to changes in the tension of their muscle coat. In a few of the viscera the

[1] Kinsella, *Mechanism of Abdominal Pain*, Sydney, 1948.

sensation is fairly accurately localized; in most it is vaguely localized or falsely referred. If the distension is moderate the sensation is one of fullness; if the distension is marked or if powerful, peristaltic movements occur, pain is experienced which may be severe in character.

(1) ŒSOPHAGUS.—If a balloon is introduced into the œsophagus, and then inflated, a sensation of fullness is felt, situated deeply beneath the sternum in the middle line; the level at which the balloon lies is located with fair accuracy. Similar accuracy of localization is generally present in carcinoma of the œsophagus. In some cases, however, the feeling of obstruction is in the upper third of the œsophagus, although the stenosis is really in the lower third; a state of spasm is then found to be present in the upper part of the œsophagus.

(2) STOMACH.—If the stomach is moderately distended with air (*e.g.* 600 c.c.), a sensation of fullness is felt deeply beneath the upper part of the anterior abdominal wall in the middle line (cf. p. 763). Greater distension (with 1500 c.c. of air) gives rise to diffuse pain which is more marked if the stomach is contracted, less marked if the stomach is relaxed (suggesting that the pain is set up in the muscle coat).

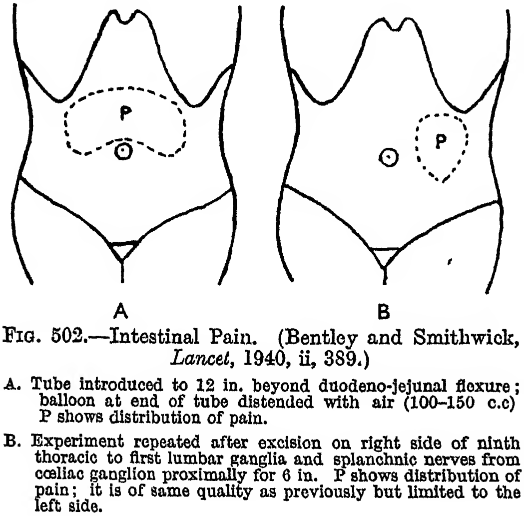


FIG. 502.—Intestinal Pain. (Bentley and Smithwick, *Lancet*, 1940, ii, 389.)

A. Tube introduced to 12 in. beyond duodeno-jejunal flexure; balloon at end of tube distended with air (100–150 c.c) P shows distribution of pain.

B. Experiment repeated after excision on right side of ninth thoracic to first lumbar ganglia and splanchnic nerves from cœliac ganglion proximally for 6 in. P shows distribution of pain; it is of same quality as previously but limited to the left side.

(3) DUODENUM.—A balloon is introduced into the duodenum in man and 10 c.c. of air injected at 10–15 mm. Hg pressure. No sensation is experienced, but reflex vasoconstriction occurs in the toes, proving that afferent impulses are set up. When 40 c.c. of air are injected at a pressure of 30 mm. Hg, more widespread reflex vasoconstriction is set up also involving the fingers. There is now also a sensation of pressure, dull and heavy in character, deep seated and poorly localized anteriorly between the xiphoid and umbilicus; at still higher pressures there is a disagreeable, painful, or griping feeling.

(4) JEJUNUM.—Distension of the jejunum about 12 in. beyond the duodeno-jejunal flexure gives rise to discomfort or pain of a similar character: Fig. 502, A shows the distribution of the pain after introducing 150 c.c. of air into the balloon. After section of the splanchnic nerves and excision of the sympathetic ganglia from Th9–L1 on the *right* side, the pain on distension was felt exclusively on the *left* side (Fig. 502, B). This observation proves that the afferents from the jejunum are conducted via sympathetic nerves into the dorsal nerve roots. Severe contraction of the intestinal wall (intestinal colic) produces pain of a character which is very generally known from personal experience. The pain is not accurately localized in the affected part as is well shown by the following case reported by Mackenzie. A loop of

small intestine was withdrawn at operation from the abdomen of a conscious subject; when vigorous peristaltic waves passed over this exposed loop the patient groaned and complained of pain which was not referred to the contracted intestine, but to the region of the umbilicus. The essential feature of the pain is its *inaccurate localization and diffuse character*.

(5) RECTUM AND BLADDER.—Distension of these organs gives rise to sensations which in part seem to come vaguely from the site of the viscus.[1]

In studies like those described above in which only the hollow viscus (stomach, duodenum, small or large intestine) and its visceral peritoneum are stimulated, the pain experienced is *not* associated with cutaneous hyperalgesia, reflex muscular rigidity or deep tenderness. It will be shown later that when as a result of disease the *parietal peritoneum* also becomes involved the new afferent fibres stimulated give rise to rigidity and tenderness as well.

Temperature Sensibility.—This is present in the œsophagus, stomach, and rectum.

(i) If very cold or hot food or drink is swallowed, a sensation of cold or heat is felt passing down the middle line under the sternum (and not in the back).

(ii) In a patient with a gastric fistula (p. 786) a balloon was inserted in the stomach and irrigated with fluid at temperatures between 18° and 40° C.; this procedure aroused no temperature response; outside these limits a temperature change of 3° C. was appreciated.

(iii) Hot or cold fluid introduced into the rectum is correctly described.

The sensibility of the stomach, appendix, and large intestine is further described below.

Stomach.—The effects of distension and heat and cold have already been considered. In the patient with a gastric fistula, light touch applied to the normal mucosa was not appreciated; light pressure (30 g. per sq. cm.) was felt as pressure, while heavy pressure (100 g. per sq. cm.) gave rise to pain; in both cases the localization was roughly correct. Severe distension gave rise to diffuse pain accompanied by nausea. Nocuous stimuli applied to the mucosa, *e.g.* pinching with forceps, faradic stimulation, application of 50–90% alcohol, N. HCl, 0·1N. NaOH, or a 3% suspension of mustard had no effect on the *normal* quiescent organ but aroused pain from the *inflamed* or the engorged mucosa. Very strong gastric contractions produced pain; the threshold was lowered when the mucosa was inflamed.

PAIN OF PEPTIC ULCER.[2]—The pain of ulcer of the stomach or duodenum is probably due mainly to the action of *acid or other irritants* on a *hypersensitive* region of the mucous membrane. The evidence is as follows:

(i) The pain tends to follow a meal after an interval which varies with the position of the ulcer; the interval is about an hour with an ulcer near the incisura and three hours or more in duodenal ulcer. Pain generally sets in when gastric acidity reaches a characteristic threshold level, *e.g.* pH 1·5 in some, or lower levels of acidity in other cases.

(ii) The ingestion of food, which decreases gastric acidity, always relieves,

[1] *Ureteric (Renal) Colic.*—When a stone passes along the ureter severe pain is experienced on the same side in the distribution of L1, often accompanied by rigidity and tenderness in the same region (p. 748). Inflammation of the ureter and periureteric connective tissue may perhaps account for the latter changes.

[2] Palmer, *Arch. int. Med.*, 1926, *38*, 694. Bonney and Pickering, *Clin. Sci.*, 1946–48, *6*, 63, 91.

and usually abolishes duodenal ulcer pain. Food often relieves gastric ulcer pain, but usually the pain has disappeared before the next meal is taken.

(iii) Vomiting, which removes acid fluid from the stomach, relieves pain generally in gastric and sometimes in duodenal ulcer.

(iv) Both varieties of peptic ulcer are relieved by alkali.

(v) In patients who are experiencing frequent bouts of *spontaneous pain*, the introduction of 200 c.c. of 0·5% HCl into the stomach when pain is absent produces pain. If the acid is introduced into the stomach when the patient is free from spontaneous pain, no pain is produced. Similar results are produced by other acids, *e.g.* H_2SO_4 or strong alkalis, *e.g.* NaOH.

In the crater of an *active* ulcer pain nerve endings are exposed and their sensitivity may be enhanced by inflammatory products or changes in the local blood supply. In these circumstances the acid which is normally secreted by the gastric mucosa, or acid or other irritants which are introduced into the stomach, may set up a sufficiently vigorous discharge from the nerve endings to give rise to a pain sensation. If the inflammation or congestion subsides the discharge set up by the less sensitive endings in response to the same stimuli may be inadequate to produce a pain sensation. Similarly during healing of the ulcer the nerve endings become covered with mucus or scab, blood clot or granulation tissue that protects them from nocuous stimuli.

Pain may however occur in peptic ulcer cases with achylia or at low levels of intragastric acidity; in these some undetermined factor must be effectively stimulating the pain nerve endings.

The *external* surface of an ulcer is generally insensitive to mechanical stimuli, *e.g.* seizure with toothed forceps.

When pain occurs in peptic ulcers, it is generally accompanied by rigidity and by *tenderness* in the epigastrium, usually over and in the recti; if the abdominal wall is anæsthetized by blocking the 5th to the 11th intercostal nerves the muscles become paralysed and the tenderness disappears. The impulses *directly* responsible for the tenderness are thus travelling in nerves coming from the abdominal wall and not from the stomach itself. But as shown in Fig. 500, impulses from the stomach acting via branching axons may be responsible for the abnormal state of the abdominal wall. The antidromic impulses on reaching the body wall release a chemical substance which enhances the excitability of the local sensory endings as explained on p. 747.

Acute Appendicitis.—In the initial stage of the disease, severe pain is felt, which is probably due to appendicular distension; it is imperfectly localized to the central part of the abdomen in the region of the umbilicus; it is unaccompanied by tenderness or rigidity. In the second stage the pain becomes localized to a restricted region on the right side and is associated with deep tenderness and rigidity. These latter signs are probably the results of impulses from the involved parietal peritoneum. It is claimed that the location of these signs varies with the position of the appendix, and depends on the region of the *parietal* peritoneum that is involved and the (somatic) afferents that are affected. The peritoneum over the posterior abdominal wall is less sensitive to mechanical stimuli than the anterior; a retrocæcal appendix in contact with this part of the parietal peritoneum may consequently give rise to little localized pain.

Large Intestine.—The crushing of a colostomy (or ileostomy) spur is

generally carried out without a local anæsthetic. If the spur is crushed piecemeal, not more than one inch at a time, the procedure is usually painless: if $3\frac{1}{2}''$ of spur are crushed at once pain frequently results. Thus in a case in which the spur consisted of healthy ascending and transverse colon the crush produced generalized deep abdominal pain, nausea, and a feeling of faintness. In another case in which the spur consisted of healthy terminal ileum and low sigmoid affected by colitis the crush produced severe continuous pain "like gripes" in the centre of the abdomen, lasting 7–8 hours and requiring two injections of morphine to make it bearable. These observations (Kinsella) again illustrate the principle that pain can be aroused in normal viscera if the stimulus is *sufficiently severe and extensive*, while slighter stimuli suffice if the viscus is made more susceptible as a result of inflammation.

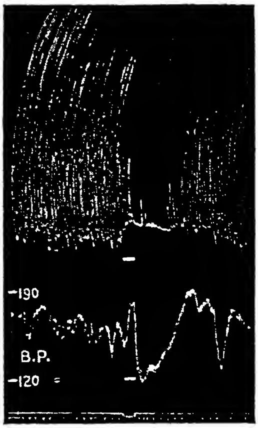


FIG. 503.—Effect of Stimulation of Central End of Vagus on Knee-jerk. (Schweitzer and Wright, *J. Physiol.*, 1937, *88*.)

Cat. Chloralose anæsthesia Records from above downwards are: knee-jerk elicited every 10 seconds; blood pressure; signal line; time in 30 sec. During the fall of the signal the central end of the vagus is stimulated. The knee-jerk is abolished. Note the "inhibitory after-discharge," *i.e.* the abolition of the knee-jerk persists for a considerable time after afferent stimulation is discontinued. The inhibition also persists while the blood pressure is returning to normal. Later on in the record the blood pressure fell spontaneously but the knee-jerk was unaffected. These last two points prove that the depression of the knee-jerk is not due to the associated circulatory changes.

Mesentery.—A pull on the mesentery readily produces pain, probably because the pain fibres from a long stretch of bowel are condensed round the blood vessels: a small stimulus can thus stimulate many afferent fibres.

Interrelationship of Autonomic and Somatic Nervous Systems.—The sharp distinction which is customarily drawn between the autonomic and somatic nervous systems, though useful for purposes of description, is to a considerable extent misleading. Afferent impulses from somatic structures may reflexly influence viscera: *e.g.* stimulation of the central end of a skin or muscle nerve may reflexly affect the heart rate or blood pressure. Conversely, afferent impulses from viscera may influence the activity of ventral horn cells (for examples cf. pp. 751, 754, 764). Both sets of results illustrate the phenomenon of *irradiation*, *i.e.* that impulses do not expend themselves exclusively in their own segment or division of the central nervous system, but tend to spread more widely to involve additional centres or regions, even some which are generally regarded as being unrelated portions of the nervous system. Several striking instances of widespread central viscero-somatic irradiation may be quoted.

(1) Stimulation of the central end of the vagus nerve reflexly inhibits the knee-jerk (Fig. 503). This effect is not due to the associated fall of blood pressure or change in breathing, but is produced by impulses which enter the medulla and pass down the spinal cord to reach the lumbar ventral horn

cells. In animals under light chloralose anæsthesia, generalized spontaneous muscular movements are common; these, too, may be reflexly inhibited by central vagus stimulation, the impulses in this case inhibiting ventral horn cells throughout the spinal cord. Even strychnine convulsions may be reflexly inhibited in the same way.[1]

(2) A rise of pressure in the carotid sinuses may produce general muscular relaxation and all the phenomena of natural sleep (Fig. 504). Bodily movements cease, the head and tail hang limply, the eyelids are partially closed, the pupils constrict, and only feeble responses to nocuous stimuli can be obtained. On lowering the pressure in the carotid sinus the animal wakes

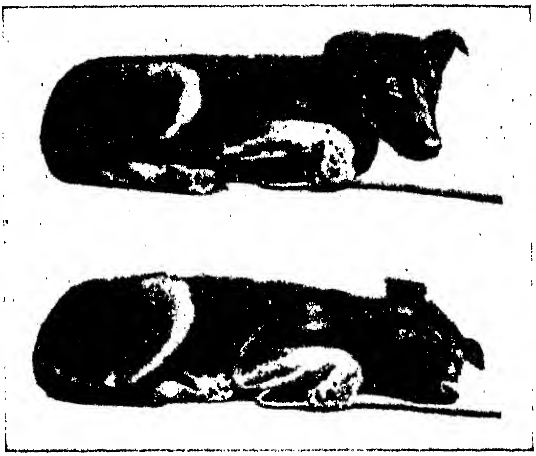


Fig. 504.—Reflex Effect of Raised Carotid Sinus Pressure on Consciousness. (After Koch, from Schweitzer, *Irradiation Autonomer Reflexe*, 1937.)

Unanæsthetized dog; right innervated carotid sinus prepared so that the pressure in it can be artificially raised from perfusion apparatus. Between A and B the sinus pressure was raised to 200 mm. Hg. In B, muscle tone disappears and the animal is in a state which resembles normal sleep.

up suddenly. Stimulation of vaso-sensory nerves may thus induce a general trance-like state (cf. p. 686).

(3) Vagus apnœa is another good instance (p. 388). The inspiratory centre (p. 384), which is a complex group of *automatically* discharging cells controlling *skeletal* muscles may be inhibited by afferent impulses coming in the vagus from a viscus, the lungs. Stimulation of the central end of the splanchnic nerves may likewise reflexly modify breathing.

(4) Where the higher levels of the brain are considered, examples of reflexes involving muscles and viscera simultaneously are numerous, *e.g.* the rage reaction of the thalamic animal (p. 664), other kinds of reflex emotional reactions or the complex responses to heat or cold. The cerebral cortex also controls visceral activity, as is seen from the results both of stimulation and extirpation experiments and of disease (p. 671).

[1] Schweitzer and Wright, *J. Physiol.*, 1937, *88*, 459.

PHYSIOLOGY OF MICTURITION[1]

Efferent Supply of Bladder.—The *efferent* fibres to the bladder come both from the *sympathetic* and the *sacral autonomic.*

The detailed anatomy of these nerves has been carefully studied in man[2] (Fig. 505). The *sympathetic* connector cells lie in the grey matter of the first and second lumbar segments of the spinal cord. The connector fibres take various routes—through the lateral sympathetic chain, and through the cœliac and superior mesentric ganglia—to form a nerve lying in front of the sacrum—the *presacral nerve.* This divides at the level of the first piece of the sacrum into two *hypogastric nerves* which end in the *hypogastric ganglia* on the lateral aspects of the rectum.[3] The *sacral autonomic* fibres arise from cells in the *second* sacral, and less constantly from the *third* sacral segments, and pass in the nervi erigentes, likewise to end in the hypogastric ganglia which serve as the excitor relay station for both sets of fibres. From the anterior border of these ganglia, sympathetic and autonomic post-ganglionic fibres arise and pass to the bladder both the body (detrusor) and internal sphincter.

The *prostatic urethra* and the *external sphincter* receive efferent *somatic* fibres from the pudic nerves.

Effects of Sympathetic Stimulation.—In animals the body of the bladder relaxes and the sphincter vesicæ contracts. Learmonth obtained the following results on stimulating the presacral (sympathetic) nerve in man: closure of the ureteric orifices, contraction of the internal sphincter, increase in the tone of the trigone and vasoconstriction in this region. In addition, there was contraction of the muscle of the seminal vesicles, the ejaculatory ducts, and the prostate, and the contained secretions were squeezed out (p. 1105). No effect was observed on the dome or lateral walls of the bladder. Intravenous injection of adrenaline in man lowers the pressure in the comfortably full bladder; it is therefore probable that the sympathetic nerves can exert some inhibitory influence on the body of the bladder in man as in lower animals.

Effects of Sacral Autonomic Stimulation.—The internal sphincter is relaxed, the detrusor is stimulated, and the bladder is emptied.

Afferent Supply of Bladder.—The afferent fibres probably take a *double* route: (*a*) along the sympathetic into the dorsal nerve roots of L1 and 2, and the lower thoracic segments; (*b*) along the sacral autonomic into the sacral dorsal nerve roots. They subserve two functions: (i) they indicate the degree of distension of the bladder; (ii) convey pain sensibility. Traction on the presacral (sympathetic) nerve gives rise to a crushing kind of pain which is felt in the bladder itself, *i.e.* it is localized surprisingly accurately and is not referred to the skin segments. The pain of bladder

[1] Fearnsides, *Brain,* 1917, *40,* 149. Head and Riddoch, *ibid.* 1917, *40,* 188–263. Denny-Brown and Robertson, *Brain,* 1933, *56,* 149.

[2] Learmonth, *Brain,* 1931, *54,* 147.

[3] In the lower animals the sympathetic relay station is in the *inferior mesenteric ganglia* (at the origin of the inferior mesenteric artery) from which hypogastric (post-ganglionic) nerves arise. In man, however, the inferior mesenteric ganglia are rarely present.

disease is diminished to a considerable extent by section of the presacral (sympathetic) nerve; it is abolished totally, however, when the sacral parasympathetic fibres are also cut.

Cortical Control.—The path of the afferent fibres within the central nervous system is not known. A higher centre for control of the bladder is described at the top of the motor area of the cerebral cortex on the medial aspect of the hemisphere in association with the centres for the perineal

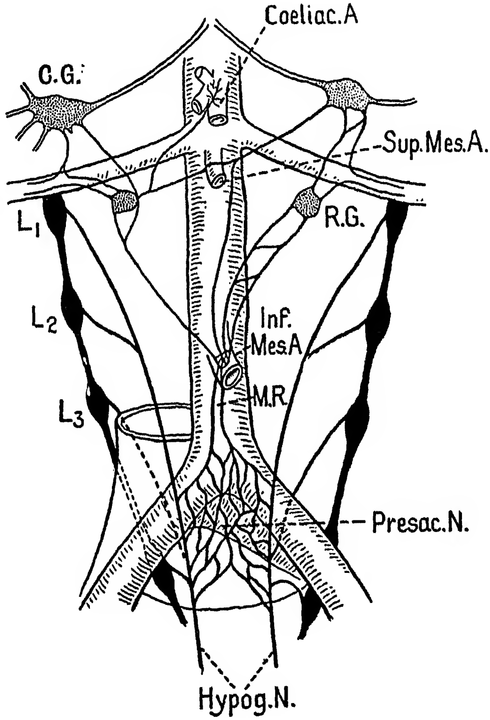


FIG. 505.—Innervation of the Bladder. (Learmonth, *Brain*, 1931.)

The diagram shows the formation of the presacral nerve (Presac. N.) from a middle root and two lateral roots. Cœliac.A., Sup.Mes.A., Inf.Mes.A. = cœliac, superior mesenteric and inferior mesenteric artery; C.G., R.G. = Cœliac ganglion, renal ganglion; M.R. = middle root of presacral nerve derived from fibres from C.G. and R.G. which come down the front of the aorta and interlace round the origin of the inferior mesenteric artery. The lateral roots come from the first to the fourth lumbar ganglia. Hypog.N. = hypogastric nerves.

structures. The path in the human spinal cord for these motor fibres from the cortex lies lateral to the pyramidal tracts intermingled with the spino-cerebellar fibres.

Postural Activity in the Bladder.—The bladder is an organ in which the phenomena of tone have been extensively studied.

(i) If fluid is injected into a bladder 24 hours *after death*, no rise of internal pressure initially takes place. The fluid serves first of all to straighten all the folds that are present; then, as more fluid is pumped in the pressure

rises with increasing rapidity for every fresh rise in bladder content. On reaching the limit of elasticity of the bladder wall the gradient of ascending internal pressure becomes very steep before the final bursting of the organ.

(ii) In the initial stage of *shock* following complete trans-section of the spinal cord, similar results are obtained, *i.e.* the bladder responds purely passively to distension with urine.

(iii) The normal bladder, however, responds differently. If fluid is introduced into the bladder through a catheter, *e.g.* in quantities of 50 c.c. at a time, the bladder pressure rises immediately on each occasion; but if an interval is allowed to elapse after each bout of filling, the pressure frequently falls to some extent, though the volume, of course, remains unchanged. These points are well brought out in Fig. 506. The intact bladder thus responds initially to distension by a contractile resistance which leads to a

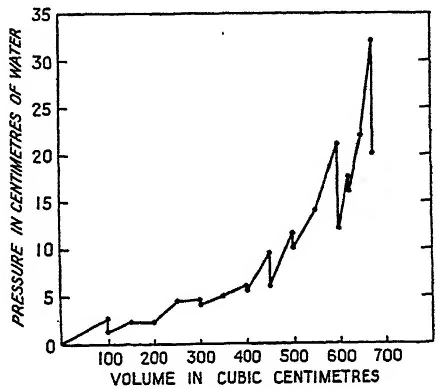


FIG. 506.—Effect of Filling Bladder on Internal Pressure.
(Denny-Brown and Robertson, *Brain*, 1933, *56*.)

Ordinate = bladder pressure in cm. H_2O ; abscissa = bladder volume in c.c.
50 c.c. of fluid were introduced into the bladder via a catheter and the resulting pressure recorded. A short interval was then allowed. The vertical drops of pressure indicate the adaptation that then takes place. Further quantities of 50 c.c. were then successively introduced and the procedure repeated Note the much more rapid rise of pressure when the bladder volume exceeds 500 c.c.

rise of pressure; subsequently the bladder wall *actively adjusts itself* to the new conditions, its fibres presumably elongate and the contents are gripped at a lower pressure than that which obtained previously. When one talks of *tone* in the bladder or other hollow viscera, one has in mind these active adjustments of the muscle coat to variations in internal volume; as a result the contents are subjected to pressures which vary to some extent, *but not directly*, with their volume. When the bladder volume exceeds 500 c.c. the pressure tends to rise more sharply and may exceed 10–15 cm. of water.

(iv) As the bladder is filled progressively there is first the gradual rise of internal pressure just described; later transient pressure waves appear which are at first small and unaccompanied by any sensation. With higher bladder volumes (*e.g.* 700 c.c.) these contractions may cause pain and are associated with a sharp rise of pressure. By an effort of the will, *i.e.* a decision to hold the urine, the contractions may pass off wholly or partially and the discomfort disappears (Fig. 507). This, however, is only temporary, for the contractions

intra-abdominal pressure compresses the bladder from without. The intra-vesical pressure rises steeply and the bladder is evacuated. Micturition can, of course, be carried out voluntarily before urgent afferent impulses have been received.

In young *children* the process is different. Postural activity in the bladder is less perfect, and quite small quantities of urine may raise the pressure sufficiently to send afferent impulses up to the cord. The higher centres are not involved and micturition occurs entirely reflexly (perhaps at the spinal level) from stimulation of the nervi erigentes.

The postural behaviour of the bladder can be influenced by a variety of circumstances. In certain forms of emotional stress, as in the anxious moments preceding an examination, the power of the bladder to elongate its fibres as its contents increase, appears to be in abeyance. The well-known result is that there is a constant desire to micturate, with the passage of only small quantities of urine on each occasion. Frequency is produced in a similar way when the bladder wall (or mucous membrane) is irritated by inflammation or stone. The nature of the disturbance in the so-called "enuresis of children" is not clear. Many authorities believe the trouble is psychical and that it is analogous to the "emotional" frequency described above.

Effects of Interference with Nervous Control of Bladder.— 1. Section of Sympathetic Supply.—In man, the immediate effects, as would be expected, are relaxation of the ureteric orifices, the trigone, and internal sphincter[1]; later, the internal sphincter may recover and close completely, though it gives way easily when a catheter is passed. After an initial and inconstant period of frequency of micturition, bladder function is re-established in a comparatively normal way.

2. Injury to Sacral Nerve Supply.—In severe cases there is complete loss of voluntary micturition; the *external* sphincter is flaccid. The bladder responds peculiarly to distension: with a small volume, *e.g.* 200 c.c., the pressure may rise rapidly to 70 cm. of water; the bladder then adapts itself quickly leading to a sharp fall of pressure to say 30 cm. of water. After a time the bladder, though deprived of its motor innervation, empties itself periodically automatically through the intervention of the local peripheral neuromuscular mechanisms. The "isolated" bladder, if otherwise healthy, can respond to adequate internal distension by contraction of the detrusor muscle; these contractions, however, are not as powerful or as well co-ordinated as normally so that micturition is incompletely performed and has to be aided by the compression effect of abdominal contraction. Large quantities of residual urine (*e.g.* 300 c.c.) may be left in the organ.

3. Injury to Afferent Supply.—The afferent impulses from the bladder may be lost or diminished from lesion of the *lumbo-sacral dorsal nerve roots*, as in tabes. The patient is unaware of the state of distension of the bladder. He can still micturate at will, but if he does not do so at regular intervals, the accumulation of urine precipitates involuntary automatic evacuation (cf. 2 *supra*); or else the pressure in the bladder may rise till the resistance of the sphincter is passively overcome and dribbling occurs (cf. p. 562).

4. Injury to Cortical Control.—Efferent control of the bladder is

[1] After *complete* sympathetic denervation, ejaculation can no longer take place, though psychical orgasm and erection occur; *sterility therefore results*

disturbed in lesions in the vicinity of the *pyramidal tracts* to the lateral side of which the fibres for the voluntary regulation of the bladder are probably found. The patient is completely conscious of all events occurring in his bladder, but voluntary control of micturition is interfered with. The patient has difficulty in initiating the act, and, further, he cannot hold his water at will: when the desire comes on he responds to it immediately. In other words, voluntary control of both motor and inhibitory fibres to the bladder is impaired. Such symptoms occur in any spinal lesions involving both pyramidal tract regions, *e.g.* disseminated sclerosis, early compression of the cord, or syringomyelia.

5. Acute Trans-section of the Cord.—If the spinal cord is completely trans-sected by some acutely acting injury, voluntary micturition is completely abolished. The activity of the detrusor muscle remains in abeyance for a long period, but sphincter tone returns very soon. At this stage the bladder responds to filling in the same way as the dead organ (p. 768) or like an elastic bag. Retention of urine is therefore complete from an early stage. If no catheter is passed, the bladder becomes increasingly overstretched. The sphincter is finally forced open by the high intravesical pressure and small quantities of urine escape at frequent intervals—a condition of "retention with overflow." Owing to the excessive stretching of the bladder wall, its nutrition suffers and it becomes very prone to infection. Cystitis may occur, and death results from the usual complications of ascending urinary infection. If the inflammation is of a lower grade, the bladder shrinks and the musculature is so damaged in consequence that it becomes incapable of ever again responding normally to internal stimuli; it contracts at irregular intervals and evacuates small amounts of urine.

If the bladder is catheterized within 24 hours of the spinal injury, fitful evacuation of urine may occur at intervals, but most of the urine has to be drawn off. As recovery occurs, the bladder responds with increasing force to distension with urine so that the pressure rises sharply; but if a constant volume is left in the bladder the pressure rapidly falls again. In the course of time, reflex micturition becomes more perfectly established. The volume of urine present in the bladder before reflex evacuation occurs shows wide individual variations. If the subject takes a deep breath his bladder, which usually empties at a content of *e.g.* 380 c.c., may empty when it contains no more than *e.g.* 80 c.c. Similar results follow when the flexor reflex is elicited or the glans penis is stimulated (p. 692). Reflex evacuation can often be produced when a catheter is in position, though it may be impossible otherwise. Presumably in such cases the coordinated inhibition of the sphincter has not yet been established, and the detrusor muscle is unable to overcome the resistance of the sphincter.

When the functional activity of the isolated region of the spinal cord becomes depressed through toxæmia, reflex bladder activity passes off in the reverse order in which it has returned; the detrusor weakens first (with retention and overflow), and finally sphincter tone is lost, and dribbling results. A similar state of affairs develops in any condition associated with extensive destruction or inflammation of the distal end of the spinal cord.

Effects of Obstruction.—Obstruction to the outflow of urine from the bladder may be due to anatomical obstacles (*e.g.* enlarged prostate or urethral stricture) or to functional derangements (*e.g.* a tonic sphincter with

a relatively weak detrusor); retention of urine consequently tends to occur, and the intravesical pressure rises till it causes expulsion of the urine. When the obstruction is mechanical, the stretching of the bladder wall acts at first as a growth stimulus resulting in hypertrophy of its fibres and increase in their expulsive power, which for a time enable the obstacle to be overcome. But when, finally, the bladder wall is overstretched, it becomes paralysed, and normal evacuation no longer takes place. At first, small rhythmical contractions may occur and account for some dribbling, but in the main the escape of urine is due to overflow from the overfilled organ. This is referred to as retention with overflow, or *passive incontinence.*

VII

DIGESTION

SECRETION OF DIGESTIVE JUICES [1]

THE SALIVA

INNERVATION OF THE SALIVARY GLANDS.—This is described on pp. 709, 714. The sympathetic fibres reach the glands along the coats of the blood vessels, the excitor cell station being in the superior cervical ganglion. The *parotid* gland receives parasympathetic fibres from the inferior salivary nucleus along the glossopharyngeal nerve; the excitor cell station is the otic ganglion which sends postganglionic fibres in the auriculo-temporal nerve to the gland. The *submaxillary* and *sublingual* glands receive parasympathetic fibres from the superior salivary nucleus through the pars intermedia and chorda tympani; the excitor relay station is in nerve cells which lie in the vicinity of the glands.

The parotid gland secretes a watery juice which is chiefly used for moistening the food. The submaxillary and sublingual glands produce a sticky juice which is rich in mucus.

Stimulation of the sympathetic is said to produce histological changes in the parotid gland and constriction of the arterioles. The parasympathetic fibres are *secretomotor* *i.e.* they cause the secretion of salts and water, and dilatation of arterioles.[2] Secretion of saliva is normally produced reflexly as a result of the taste, sight, thought, or smell of food.[3]

FUNCTIONS OF SALIVA.—These are mechanical and digestive. Saliva aids mastication and swallowing, it facilitates speech, and helps to dilute irritants and to cool excessively hot foods. The unimportant digestive action of ptyalin takes place mainly in the stomach and is concerned with the initial stages of the hydrolysis of starch (p. 835).

The relation of dry mouth to *thirst* is considered on p. 68.

[1] Pavlov, *Work of Digestive Glands*, 2nd edn., London, 1910. Babkin, *Secretory Mechanism of Digestive Glands*, 2nd. edn., New York, 1949. Grossman (Hormones of Alimentary Canal), *Physiol. Rev.*, 1950, *30*, 33.

[2] For discussion of the changes which take place in these glands during activity, see p. 20; see pp. 675 *et seq.* for use of salivary secretion in the study of conditioned reflexes.

[3] ŒSOPHAGO-SALIVARY REFLEX.—If the œsophagus is cut and its distal portion is stimulated, an abundant flow of saliva results. This reflex depends on the integrity of the vagi. The importance of this reflex is that it enables any bolus which is held up in the œsophagus to be washed on. Peristalsis cannot be initiated in the œsophagus by means of mechanical stimulation; but when the reflexly secreted saliva is swallowed, it sets up the processes of deglutition which result in coordinated contraction of the œsophagus, and the bolus is moved on. In *carcinoma* of the œsophagus, salivation, reflexly produced in the way described, is a common symptom.

GASTRO-SALIVARY REFLEX.—When food is introduced into the stomach of a dog with precautions to avoid psychic impressions, a flow of saliva occurs after an interval of 20 minutes.

SECRETIONS AND FUNCTIONS OF GASTRIC JUICE

Gastric Mucous Membrane.—Three types of glands are found in the gastric mucosa (Figs. 113, 508)[1]:

(1) The *cardiac glands* occupy a zone a few millimetres wide near the cardiac orifice of the stomach; they are mucus secreting glands resembling those of the pylorus. Nothing is known about their function or control.

(2) The *pyloric glands* commence at the incisura angularis on the lesser curvature; on the greater curvature they generally begin at a point considerably nearer to the pylorus. The ducts are long and the gland alveoli short; the latter resemble Brunner's glands of the duodenum and secrete an *alkaline* fluid which is rich in *mucus* and has considerable neutralizing power but is poor in chloride and digestive enzymes.

Peripheral vagus stimulation (and in some species sympathetic stimulation) leads to secretory activity in this region; little is known about the general regulation of pyloric secretion.

(3) *Main gastric glands.*—The rest of the gastric mucosa (in the fundus, body, and part of the pyloric region) contains what may be called the main gastric glands (often referred to misleadingly as fundic glands). The ducts are short and the gland alveoli are long; the latter are lined by two types of cells: the *chief* or *peptic* which secrete *pepsin* (from an inactive precursor, *pepsinogen*), and the ovoid *parietal* or *oxyntic* cells which secrete HCl (Fig. 508).

(4) The surface lining of the stomach consists of columnar epithelial cells between which lie many mucus secreting goblet cells; there are also many mucous cells in the necks of the glands.

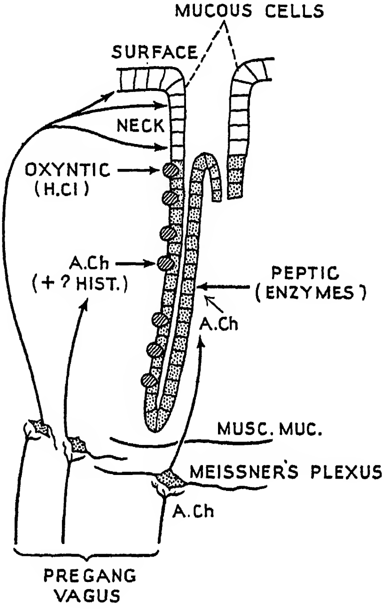


FIG. 508.—Diagram of Nerve Supply of Gastric Mucosa and of Chemical Transmitters Concerned.

Pregang. Vagus: preganglionic fibres of vagus end in Meissner's plexus. Transmitter is A.Ch (acetylcholine).
Postganglionic fibres from Meissner's plexus end in gland cells. Transmitter to peptic cells is acetylcholine; to oxyntic cells it is acetylcholine (+histamine ?).

The main gastric glands secrete the fluid generally called the gastric juice. This contains (Fig. 59):

(i) A strongly acid solution from the oxyntic cells containing 0·4–0·5% HCl (*i.e.* stronger than 0·1N which is 0·35%) and varying little in composition[2]; it is essential for peptic digestion and has some antiseptic action.[3]

[1] The arrangement of the *muscle fibres* of the stomach is described on p. 807.

[2] Experiments on isolated gastric mucosa (frog) show that the oxyntic cells can secrete HCl at a concentration of 0·16N (pH=1·4).

[3] *Bacteria* up to 100,000 organisms per c.c. may also be found, chiefly bacillus coli, staphylococci, and non-hæmolytic streptococci. The *bactericidal power* of the gastric juice probably prevents virulent organisms (*e.g.* hæmolytic streptococci) from entering the duodenum alive.

(ii) *Pepsin* (and perhaps rennin and lipase) from the chief cells. Pepsin (in the presence of HCl), breaks proteins down to peptones (p. 878). Rennin curdles milk; it acts on the caseinogen of milk and, in the presence of ionic calcium, insoluble calcium caseinate is precipitated.

(iii) *Intrinsic factor* (in man) (p. 201).

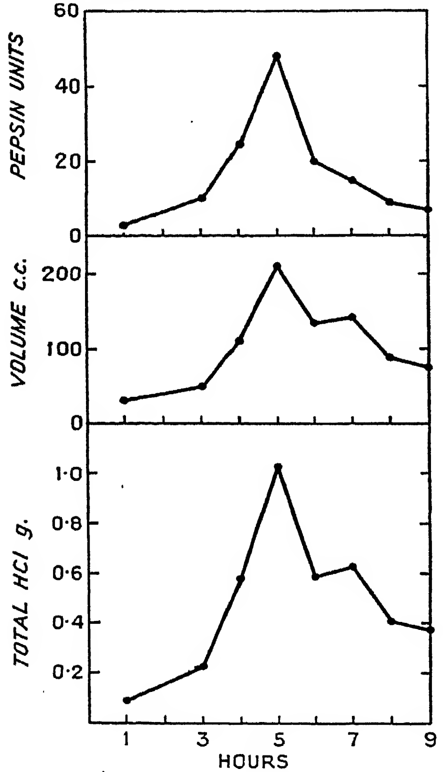


FIG. 509.—Effect of Continued Vagus Stimulation on Gastric Secretion. (Bowie and Vineberg, *Quart. J. exp. Physiol.*, 1935, 25.)

Dog: peripheral vagus stimulated for 9 hours.
The ordinates from above downwards are:
(i) pepsin in arbitrary units,
(ii) volume of gastric juice in c.c.,
(iii) total HCl in g., secreted per hour.
Note the increase in all three values.

(iv) *Mucin*, both dissolved and visible, from the superficial and neck cells.

Peripheral or reflex vagal stimulation causes simultaneous secretion of the enzymes, HCl, and mucus from the main gastric gland area (Fig. 509); these constituents are not necessarily always secreted in the same proportions. The secretion is accompanied by vasodilatation; the mucosa becomes rose red in colour.

The gastric mucosa also releases several hormones into the blood stream. The best known are *gastrin* and *enterogastrone*.

The chemical transmitter of the vagus to the stomach (as elsewhere) is acetylcholine (p. 719, Fig. 508). After meals or vagus stimulation normal gastric juice also contains *histamine* which is unrelated in amount to the histamine concentration in the blood.[1] The physiological significance of this histamine content is uncertain, but it has been suggested that in the case of the oxyntic cells of the stomach two chemical mediators are successively involved: stimulation of the vagus first releases acetylcholine; this in turn releases histamine which penetrates into the oxyntic cells to cause a discharge of HCl (Fig. 508); some of this histamine secondarily escapes into the gastric juice.

Collection of Pure Gastric Juice.—In animals (dog) the following procedure is employed. A special stomach pouch is prepared as described by Pavlov (a *Pavlov pouch* (Fig. 510)). A small pouch of the stomach is separated from the main body of the organ by a double layer of mucous membrane; the open end is brought up to the surface of the body. The nervous and vascular connections of the pouch are left intact. Pure gastric juice unmixed

[1] Code, Hallenbeck, and Gregory, *Amer. J. Physiol.*, 1947, *151*, 593.

with food can be obtained from the pouch while the digestive processes are proceeding in the main stomach. The vagus nerve is exposed in the neck under anæsthesia and divided. A few days are allowed to elapse for the inhibitory fibres to the heart to degenerate; for some unknown reason the vagal fibres to the stomach (and pancreas (cf. p. 790)) survive longer than those to the heart. The peripheral end of the vagus may then be stimulated in the unanæsthetized animal; a flow of gastric juice results after a short latent period, proving that the vagus is the secretory nerve to the stomach (Fig. 509).

Collection of Gastric Juice in Man.—This is usually aspirated from the stomach after swallowing a fine rubber tube (cf. p. 781). In cases of *gastric fistula* (resulting from injury or operation), pure juice may be directly obtained and the mucosa can be readily inspected (p. 786).

Regulation of Secretion of Gastric Juice.[1]—There are three phases of gastric secretion. The first is the *cephalic* or *nervous* phase which is due to reflex stimulation of the *vagus* nerves; the juice so formed is called psychic or appetite juice. This is followed by the *chemical* phase, which is subdivided into a *gastric* and an *intestinal* phase. In the chemical phase, *hormones* are absorbed into the blood stream, reach the stomach again via the circulation, and modify its secretory activity; the hormones include the excitatory agent *gastrin*, and the inhibitory agent *enterogastrone*. Absorbed products of digestion may act similarly. The regulation of gastric secretion is most readily investigated by means of animal experiments which are described below; human observations will in the main be considered later (p. 781).

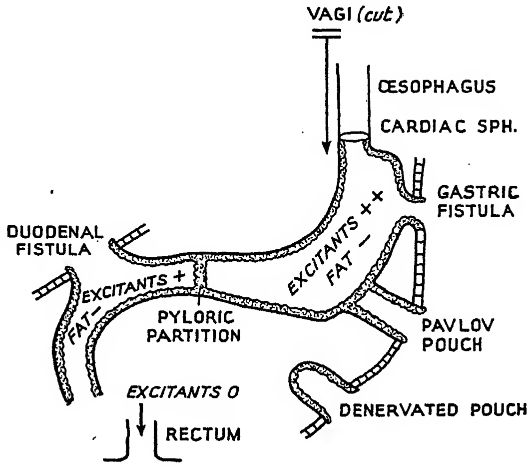


Fig. 510.—Diagram illustrating Experimental Analysis of Humoral Factors regulating Gastric Secretion. The pyloric partition is at the pyloric sphincter.

1. Nervous Phase. Appetite Juice.—If the vagi are intact, the sight, smell, or taste of food reflexly produces a flow of gastric juice. The secretion thus obtained is the *psychic* or *appetite* juice. Taste is the most effective stimulus and sets up an inborn reflex via the brain stem; the sight and smell of food set up secretion by means of an acquired or conditioned reflex (via the cerebral cortex (p. 678)). Chewing indifferent substances not related to food is ineffective. Appetite juice is well demonstrated in animals by the method of *sham feeding*. The œsophagus is exposed in the neck and divided, and the two cut ends are brought up to the surface. Sham feeding or drinking can now be carried out, *i.e.* food is given which is masticated and swallowed, but drops out through the upper end of the cut œsophagus. Appetite juice

[1] Kahlson, *Brit. med. J.*, 1948, ii, 1091.

(in the dog) in response to sham feeding sets in after a latent period of 5–10 minutes; the volume secreted reaches its peak after 1 hour and persists for 3–6 hours. The juice is highly acid and has a high pepsin concentration.

Appetite juice is not essential to health. When it has been abolished by section of the vagi, digestion is not impaired. It is of value, however, in initiating the digestion of the food; some of the products of such digestion by releasing gastrin and perhaps in other ways produce a further flow of gastric juice (*infra*).

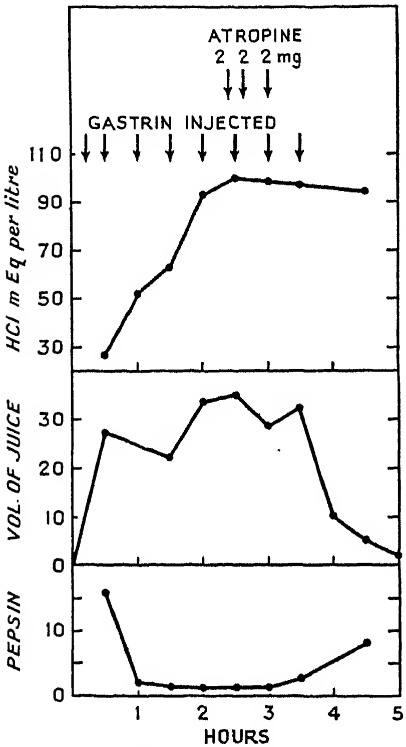


FIG. 511.—Effect of Gastrin on Gastric Secretion. (Komarov, *Rev. Canad. Biol.*, 1943, *1*, 380.)

Cat under choralose; cannulate entire stomach. The gastrin preparation was histamine-free and had no effect on blood pressure. Each arrow represents the injection of 20 mg. of gastrin.
Records from above downwards are:
(i) HCl in m.Eq./L.,
(ii) volume of gastric juice, in c.c.,
(iii) pepsin in arbitrary units, secreted per half hour.
Gastrin stimulates secretion of gastric fluid and acid, but not of pepsin. The response is not annulled by atropine. (cf. Fig. 522)

Appetite juice may be secreted *in man* in response to the sight or smell of appetising food, or when a meal is chewed and immediately spat out; the secretion is usually of brief duration but is accompanied (as in animals) by an increased blood flow through the gastric mucosa and increased motility. The effects of fear and anxiety are considered on p. 786.

2. Chemical Phase. Gastric Response to Food.—If certain foodstuffs are introduced into the stomach directly via a gastric fistula (Fig. 510) a secretion of gastric juice is obtained after a latent period of 30–60 minutes. The volume of juice reaches its peak in 2 hours and secretion continues for some hours afterwards; the juice has a high acid content but its pepsin content is lower than that of vagus juice. Under natural conditions the appetite juice is followed by the chemical response to the presence of food in the stomach.

MECHANISM OF CHEMICAL PHASE.—The following substances produce a secretory response: (i) meat extracts; (ii) certain foodstuffs, *e.g.* bread, egg white, meat, milk, and their digestion products; (iii) liver extracts. Distension of the stomach also stimulates gastric secretion. All these procedures produce their effects by releasing gastrin into the blood stream.

The chemical phase is not due to reflex secretion via the vagi as it persists after cutting both these nerves. It is not due to a local reflex through the intrinsic gastric plexus (*i.e.* Meissner's plexus); thus when a suitable food is introduced into the main stomach, a secretory response is obtained from a completely separated and *denervated stomach pouch* (Fig. 510). This last

experiment also proves that the chemical excitants do not stimulate the gastric glands by a direct action. It may be concluded therefore that a hormone (gastrin) is absorbed into the veins, enters the general circulation and reaches the gastric glands via their arterial blood supply.

The structures which form and release gastrin into the circulation have not been identified; their site can be deduced by indirect methods.

(i) Extracts of the pyloric region, the adjacent region of the body of the stomach ("intermediate zone") and the duodenum contain gastrin.

(ii) The stomach is shut off from the duodenum by means of a double fold of mucous membrane at the pylorus (Fig. 510). The chemical excitants are then found to act most strongly when introduced into the stomach, but they are also effective, though to a less extent, when introduced directly into the duodenum; they are inactive when introduced into the rectum. The chemical response can thus be divided into a *gastric* and an *intestinal* phase.

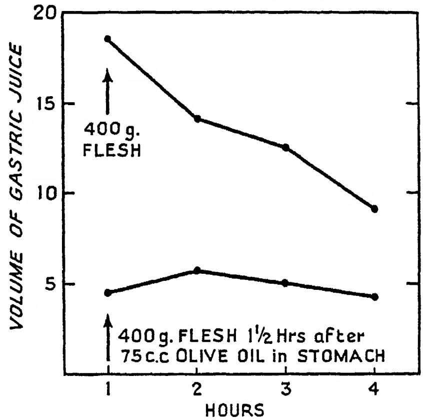


FIG. 512.—Inhibitory Effect of Fat on Gastric Secretory Response to Food. (Data from Pavlov, *Work of Digestive Glands.*)

Experiments on dog with Pavlov pouch. Time in hours after ingesting meat. Vertical axis represents volume of gastric juice in c.c. secreted per hour by the pouch.
Upper curve: after eating 400 g. of meat.
Lower curve: same meal eaten 1·5 hr. after introduction of olive oil into the stomach. Note marked inhibitory effect of the oil on the gastric response.

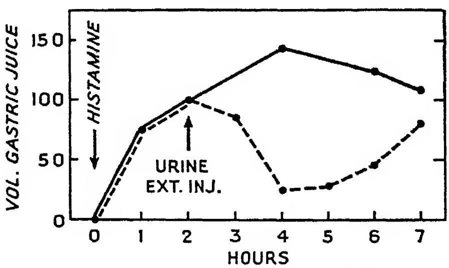


FIG. 513.—Inhibitory Action of Urogastrone on Gastric Secretion. (Friedman and Sandweiss, *Amer. J. dig. Dis.*, 1941, *8.*)

Experiment on dog: gastric juice aspirated.
Upper continuous line=control=response to injections of 0·1 mg./kg. of histamine subcutaneously hourly. Lower broken line=response to same injections, but at arrow, urine extract (urogastrone) was injected. Note inhibition of secretion.
Vertical axis: volume of gastric juice secreted per hour; 100=100%=volume secreted during second hour.

(iii) An *isolated denervated* pyloric pouch is prepared. If it is mechanically distended the main stomach responds by secreting juice. This experiment proves that the pylorus releases gastrin; this region is probably the main site of gastrin formation.

GASTRIN.—Gastrin is a protein of low molecular weight, closely resembling secretin (p. 792) in its chemical properties; it is free of histamine and has no effect on the blood pressure. It stimulates the stomach to secrete a juice which is poor in pepsin but rich in acid (Fig. 511); it is believed to act specifically on the oxyntic cells. It has no action on the salivary glands, pancreas, or

liver or on gastric motility. It is effective only when injected intravenously; intramuscular injections induce a very scanty secretion.[1]

A *fall of blood sugar* (Fig. 515) or the intravenous injection of certain amino-acids (glycine, alanine, glutamic acid) increases gastric secretion of both HCl and pepsin by stimulating the *vagus nucleus* in the medulla.[2] A rise in blood sugar inhibits gastric secretion (Fig 516).

Inhibitory Agents : Enterogastrone, Urogastrone.—The introduction of *fat* into the stomach inhibits gastric secretion of both HCl and pepsin in response to a meal (Fig. 512) ; a much greater inhibitory effect follows the introduction of fat directly into the duodenum. Fat also inhibits the gastric secretory response to sham feeding, injection of histamine (Fig. 515) or insulin (p. 785). As the introduction of fat into the duodenum inhibits the secretion of a transplanted *denervated* gastric pouch, it must act through a humoral mechanism. A gastric depressant agent has been detected in the blood of animals fed on fat. A potent inhibitory extract has been isolated from the *intestinal* mucosa and called *enterogastrone* ; it is thought to be responsible for the inhibitory action of fat. Ingestion of fat and injection of enterogastrone inhibit gastric motility too. A substance similar to, but not identical with, enterogastrone has been extracted from the urine in man and called *urogastrone.* Fig. 513 shows the inhibitory action of urogastrone on the gastric response to histamine.[3]

Gastric Response to Different Foodstuffs.—As might be expected from the complex mechanisms involved there is a fairly characteristic gastric response for each kind of food in respect of volume and composition of secretion.

Rôle of Hypothalamus.—Experimental lesions of the hypothalamus in animals are often associated with gastric hæmorrhage, erosions or even perforation. The probable clinical significance of these results

[1] Gastrin probably does not account completely for the chemical phase of gastric secretion ; when the chemical excitants are introduced into the stomach they produce a gastric juice which has a *higher pepsin content* than that induced by gastrin. Crude pyloric extracts have been prepared which contain another principle which also increases pepsin secretion ; if this work is confirmed this substance might be called " gastrozymin " (cf. pancreozymin p. 793). According to Uvnäs (*Acta physiol. scand.*, 1942, *4*, Suppl. xiii) gastrin (or some associated pyloric hormone) is also linked up with the gastric *vagal* secretory mechanism. The vagal fibres which supply the pylorus are alleged to stimulate the release of gastrin ; *i.e.* though gastrin is released mainly by chemical factors it is also released to a minor extent by nervous factors. It is claimed that *circulating* gastrin must be present if the vagus is to exert its secretory action on the main gastric glands. If the pylorus is excised, deprived of its blood supply, or poisoned with cocaine, stimulation of the vagus in the fasting animal fails to elicit the usual profuse flow of an acid- and pepsin-rich juice from the main gastric glands ; but if gastrin is infused into these animals (in concentrations which in themselves are ineffective) concurrently with vagal stimulation, the normal response characteristic of the intact animal is obtained. It is argued that gastrin is part of the transmission mechanisms at the vagal terminals in the stomach. It would follow from this work that vagotomy in addition to abolishing the nervous control might also depress the responses of the stomach to chemical stimuli. Similarly pylorectomy by removing the chief gastrin-bearing area might depress the efficacy of the vagal secretory control. [There may be a similar relationship between the action of the vagi and of *Secretin* on the pancreas.]

[2] It is surprising that glucogenic amino-acids should have the same effect as a *fall* of blood sugar.

[3] It is claimed that though uncooked fat (*e.g.* olive oil) inhibits gastric secretion, certain *cooked* fats stimulate secretion. The point requires further study.

(especially in relation to the production of peptic ulcers) is considered on pp. 786 *et seq.*

Gastric Secretion in Man.—FASTING STOMACH.—Beaumont, in his classical observations on Alexis St. Martin, who had a gastric fistula following a grave injury to the abdominal wall, found that the fasting stomach was empty and contracted; he applied his tongue to the gastric mucosa but noted no acid taste. Carlson, on the other hand, found in healthy individuals with gastric fistulæ that the fasting stomach secretes gastric juice continuously at a rate of 10–60 c.c. per hour. The discrepancy may be due partly to individual peculiarities and partly to the fact that Carlson's patients were studied with a rubber tube lying in the stomach; Beaumont observed that a gum elastic tube is a strong mechanical excitant to gastric secretion in man (this incidentally is contrary to Pavlov's experience in the dog). Clinical tests of gastric secretory function all involve the presence of a tube in the stomach.

It is claimed that patients with duodenal ulcer secrete during the night a larger volume of gastric juice, which is also richer in acid than in normal people (cf. p. 787).

Clinical Investigation.—1. Fractional Test Meal.—The course of gastric secretion in man can be followed by means of the fractional test meal. A fine flexible rubber catheter is swallowed first thing in the morning, no food having been taken since the previous evening. The *resting contents* of the stomach are aspirated with a syringe. The stomach is washed out with distilled water which is withdrawn. A pint of thin gruel is swallowed and 10 c.c. of gastric contents are aspirated at intervals of 15 minutes until the stomach is empty. Each sample is examined to determine the free HCl and total acidity. The results are plotted in the form of a curve (Fig. 514).[1]

A careful distinction must always be drawn between gastric *contents* and pure gastric *juice*. The contents of the stomach consist of varying proportions of acid juice secreted by the main gastric glands, mucus and alkali secreted by the pyloric glands, swallowed saliva, swallowed food, and pancreatic and intestinal juices which have regurgitated through the pylorus.

(i) *Resting Gastric Contents.*—After a night's fast, the stomach contains a varying volume of fluid, 30 c.c. on the average; free HCl can usually be demonstrated (so-called *fasting secretion* (cf. *supra*)). The fluid consists of gastric juice, saliva, and mucus, and may be stained with bile which has

[1] The technique is as follows: To 5 c.c. of gastric contents a few drops of Töpfer's reagent (dimethyl-amino-azo-benzene) are added. A bright red colour indicates the presence of *free HCl.* Titrate with 0·1N NaOH till the solution becomes lemon-yellow. The results are always expressed in terms of the volume of 0·1N NaOH which neutralizes *100 c.c.* of gastric contents (or the equal number of c.c. of 0·1N HCl which have been neutralized). Add a drop of phenolphthalein, and run in more NaOH till the solution goes pink; this gives the *total acidity.* The difference between the free HCl and total acidity is a measure of the HCl which is organically combined with mucin and protein (and of any abnormal acids like lactic acid, which may be present). The total acidity usually exceeds the free HCl by about 10 c.c. of 0·1N HCl.

In addition, the *total chloride* content may be determined; this represents the sum of the free and combined HCl and the *inorganic chloride*; the last is mainly produced by neutralization of the gastric HCl by alkali.

All samples are tested with iodine for the presence of *starch,* to determine the *emptying rate* of the stomach. When the starch reaction is no longer obtained (usually after 2½ or 3 hours), it is assumed that all the gruel has left the stomach. A careful note is also made of the appearance of *bile* and *mucus* (or abnormal constituents, *e.g. blood*).

regurgitated from the duodenum. If the stomach has been secreting rapidly, the stomach contents may sometimes approach in composition that of pure gastric juice.

(ii) Ingestion of the pint of gruel usually lowers the free acid almost to zero level, owing to (*a*) dilution of the gastric juice; (*b*) its neutralization by swallowed saliva; and (*c*) its combination with the organic substances in the meal.

(iii) The acidity then usually slowly rises, and after 1½ hours reaches 0·1% HCl (equivalent to 30 c.c. of 0·1N NaOH per 100 c.c. of gastric contents) on the average. The presence of the food and water in the stomach has thus called forth a second flow of gastric juice (chemical phase); as this secretion is added to the stomach contents, the acidity slowly rises (Fig. 514). Subsequently the curve of acidity may go on rising or, on the contrary, may fall rapidly after the initial rise; some cases show very little rise even at first.

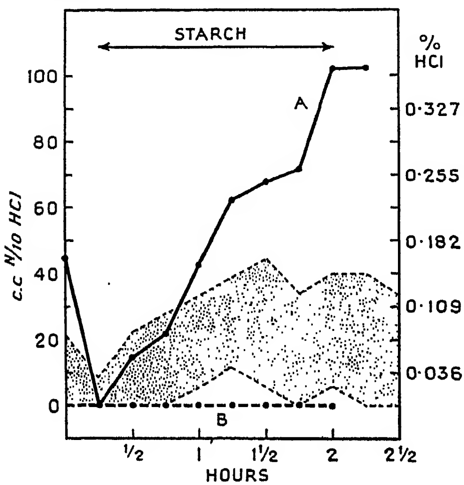


FIG. 514.—Fractional Test Meal. (Bennett and Ryle.)

Ordinate: c.c. of N/10 HCl (left) or g. of HCl (right) in 100 c.c. of aspirated gastric *contents*.
Shaded area represents the limits of free HCl in 80 per cent. of healthy males after a test meal. A= "Climbing" type of curve; B=Achylia gastrica; The emptying time of stomach is measured by the disappearance of starch (in this Fig. after 2 hours).

FACTORS INFLUENCING ACIDITY CURVES.—It is hardly surprising that the acid curves obtained show such wide variations in normal subjects. The following factors probably play a part:

(i) The *quantity* of acid secreted in response to the meal: in some people the stomach secretes a large, in others a small, volume of juice. (The HCl concentration in the juice *as secreted* is probably fairly constant, *i.e.* over 0·1N, about 0·4–0·5%.) Other things being equal, the larger the amount of HCl which is secreted and thus added to the gastric contents the higher will be the concentration of acid attained. A small percentage of normal subjects secrete no HCl at all (curve B, Fig. 514).

(ii) The *rate* at which the gruel *leaves* the stomach: if the gruel leaves rapidly and the stomach continues to secrete actively then the gastric contents approximate more and more closely to pure gastric juice with its high acidity; the result is a "climbing" curve like A in Fig. 514.

The above factors influence the rate and extent of the *rise* of intragastric acidity; they obviously cannot produce the *fall* of acidity which usually occurs during the latter half of the test. This decline can only be brought about by *neutralization* of the acid gastric contents by alkali. This alkali is derived to some extent from regurgitated intestinal juices but comes mainly from the alkaline secretion of the pyloric glands.

(iii) *Rôle of Regurgitation of Intestinal Contents.*—Intestinal contents sometimes regurgitate into the stomach (p. 810); this is proved by the not un-

common occurrence of bile in some of the aspirated samples (usually the later ones). The upper duodenum contains *alkaline* juices secreted by the liver, pancreas, and the intestinal mucosa. When this alkaline fluid enters the stomach it neutralizes the gastric HCl and so lowers gastric acidity. A "flat" test-meal curve may be due to free regurgitation, while a "climbing" curve is favoured by absence of regurgitation.

(iv) *Rôle of Gastric Mucus and Alkali.*—The surface epithelial cells of the stomach, the neck cells of all the glands, and the alveoli of the pyloric (and cardiac) glands secrete mucus; the pyloric glands also secrete an alkaline fluid. Mucus can neutralize considerable amounts of HCl; thus if the juice secreted by a Pavlov pouch is allowed to remain in contact with the surface mucus for only 5 minutes, the acidity may fall by 25%. The importance of *intragastric* neutralization of the gastric contents in regulating intragastric acidity is shown by this fact: if the duodenal contents are steadily aspirated during the progress of a test meal the character of the test-meal curve may be entirely unaffected and the decline in acidity may occur in the usual way.[1]

The clinical value of the fractional test meal is limited; the *character of the curve obtained has little diagnostic value.* Some light is thrown by this means on the emptying rate of the stomach; but information about gastric motility is best obtained by radiographic examination. *Pyloric stenosis* can, however, be diagnosed from the presence of a large volume of resting (fasting) gastric contents which may be dark in colour and foul as a result of putrefactive change. The presence of blood is strong evidence of an organic lesion, *e.g.* ulcer or cancer.

2. Histamine Test of Gastric Function.—Histamine is a powerful stimulant to the secretion of gastric juice; it acts mainly on the oxyntic cells causing secretion of an increased volume of highly acid gastric juice with a low pepsin content (Fig. 515). The test is carried out as follows: the resting gastric contents are first aspirated with a stomach tube and the stomach washed out. 0·5 mg. of histamine is injected subcutaneously.[2] The secretion is continuously aspirated; the volume and acidity of each 10-minute sample are determined.

The test is not unpleasant, though a quarter of the subjects develop headache. Intense salivation also takes place and the subject must be careful not to swallow the saliva, as that will invalidate the gastric analyses. Following the injection, the gastric acidity rises steadily to a maximum and then falls off. Attention is paid to the following points (*mean* standards are indicated in brackets):

(i) Highest level of acidity obtained (102 c.c. of 0·1N HCl).

(ii) Volume of 0·1N acid secreted in 1 hour (182 c.c.).

(iii) Volume of juice secreted in 1 hour (200 c.c.). For (ii) and (iii) the data for the specimens collected from the 10th to the 50th minute are used (*i.e.* for 40 minutes in all), and multiplied by 1·5.

An injection of histamine tests almost specifically the functional state of the oxyntic cells; it may be regarded as mimicking the action of gastrin. As the gastric juice is secreted rapidly and aspirated continuously, neutralizing factors are reduced to a minimum; the peak of acidity noted after histamine

[1] Baird, Campbell, and Hern, *Guy's Hospital Rep.*, 1924, *74*, 23.

[2] It is very important to note that if the initial blood pressure is under 110 mm. Hg the test may be dangerous.

injection probably represents the acid concentration of pure human gastric juice. The test gives a quantitative idea of the activity of the gastric mucous membrane. Histamine produces marked congestion of the mucous membrane in man.

In a small proportion of normal subjects (perhaps about 2%) the gastric

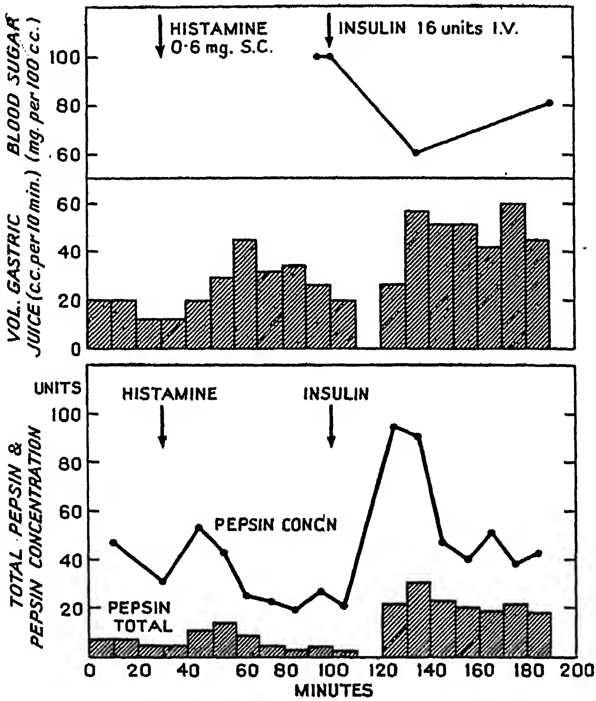


FIG. 515.—Comparison of Action of Histamine and Insulin on Gastric Secretion in Man. (Ihre, *Acta med. scand.*, 1938, *95*.)

Records from above downwards: Blood sugar in mg. per 100 c.c.; volume of gastric juice (in c.c. per 10 min.); pepsin concentration in each sample and total pepsin content of each sample (the last two in arbitrary units).
At first vertical pair of arrows: inject 0·6 mg. of histamine subcutaneously.
At second vertical pair of arrows: inject 16 units of insulin *intravenously*.
Note that insulin produces a much greater total pepsin secretion and pepsin concentration; the peak of the insulin response coincides with the maximum fall of blood sugar (down to 60 mg–%.)

mucosa is undeveloped and secretes no HCl; the condition is called *achylia gastrica*. A similar achylia (accompanied by extensive destructive changes in the gastric glands) is a regular accompaniment of *pernicious anæmia* (p. 197). Achylia of a transient character may result from inflammatory conditions of the gastric mucosa. The absence of free acid in the fractional test meal is not conclusive proof of absence of HCl secretion as it may equally well be due to excessive neutralization; but a negative response to histamine, especially after repeated injections, is satisfactory proof of achylia.[1]

[1] See Watkinson and James, *Clin. Sci.*, 1951, *10*, 255, for possible fallacies.

Patients with a hypertrophic gastric mucosa respond more markedly than normal subjects to histamine.

In *cancer* of the stomach, free HCl is often absent from the gastric contents. This is probably due to the inflamed condition of the mucosa. If the interior of the stomach is washed to remove adherent mucus, evidence of gastric secretion can usually be obtained, especially in response to a histamine injection.

3. **Insulin Test of Gastric Function.**—The injection of insulin (7 units *subcutaneously*) is followed in man by a marked secretion of gastric juice resembling that produced by peripheral vagal stimulation, *i.e.* it is rich both in HCl and pepsin content. The response sets in after a latent period of about 40 minutes and coincides with a considerable fall of the blood sugar, *e.g.* to 70 mg–%, and the usual early signs of hypoglycæmia (p. 915) such as dizziness or drowsiness. Fig. 515 shows a very striking response to the injec-

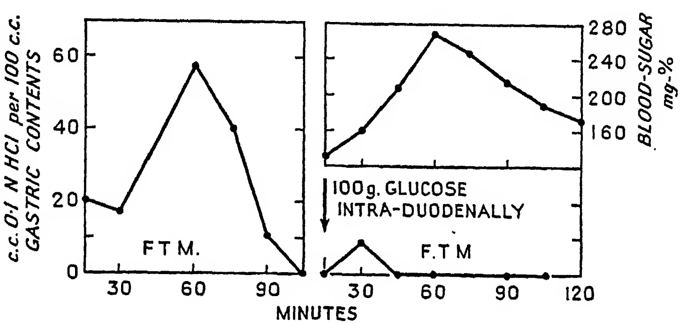


FIG. 516.—Inhibitory Effect of Raised Blood Sugar on Gastric Secretion in Man. (Glaessner, *Amer. J. dig. Dis.*, 1943, *10*.)

Left-hand curve : F.T.M. = normal response to gruel meal.
Right-hand curves : upper : blood sugar in mg–% ; lower : F.T.M. = gruel meal after administration (at arrow) of 100 g. of glucose intraduodenally.
Note that acid secretory response is almost completely annulled ; blood sugar rises to 270 mg–%.

tion of 16 units of insulin *intravenously*. The gastric response is annulled when glucose administration is combined with the insulin injection to prevent the fall of the blood sugar ; it is clear that insulin does not act directly on the gastric secretory mechanisms but produces its effects solely by virtue of the resulting hypoglycæmia. In severe diabetic subjects even large doses of insulin fail to stimulate gastric secretion unless the blood sugar is reduced to a subnormal level. Appropriate cross-circulation experiments demonstrate that the hypoglycæmia acts on the central nervous system and causes increased activity of the vagus secretory supply to the stomach. The gastric response to hypoglycæmia is abolished, as expected, by double vagotomy (performed in man for peptic ulcer) (p. 789) or by suitable doses of atropine. Insulin injection (like vagus stimulation in animals) markedly increases gastric vascularity in man (cf. p. 776).

Effect of Raised Blood Sugar:—A marked rise of the blood sugar in man abolishes the secretion of acid by the stomach in response to a gruel test-meal. This is well demonstrated in Fig. 516 ; following the introduction of 100 g.

of glucose intraduodenally which raised the blood sugar to 270 mg–% a gruel test-meal produced no acid secretion whatever.

4. **Combined Insulin-Histamine Test.**—Insulin (*e.g.* 7 units) is injected first, followed 20 minutes later by an injection of histamine (0·5 mg.); the gastric response sets in about 20 minutes later still. This technique subjects the stomach simultaneously to the stimulating action of the vagus nerve and of the peripherally acting chemical agent. The *maximum* secretory powers of the gastric mucosa can thus be determined. The volume of juice secreted during the period 40–60 minutes after the insulin injection is measured; on the whole it varies with the state of the gastric mucosa as seen gastroscopically. With a normal-looking mucosa the volume is about 2·5 c.c. per minute and the maximum acidity is equivalent to about 100 c.c. of 0·1N HCl (Fig. 517). When the mucosa is swollen and hyperæmic much greater reactions are obtained; the volume may be as high as 6·4 c.c. per minute and the maximum acidity may exceed 0·16N HCl. Fig 517 also shows (as would be expected) that the combined insulin - histamine stimulus produces a far higher level of acidity of the gastric contents than does the fractional test meal.

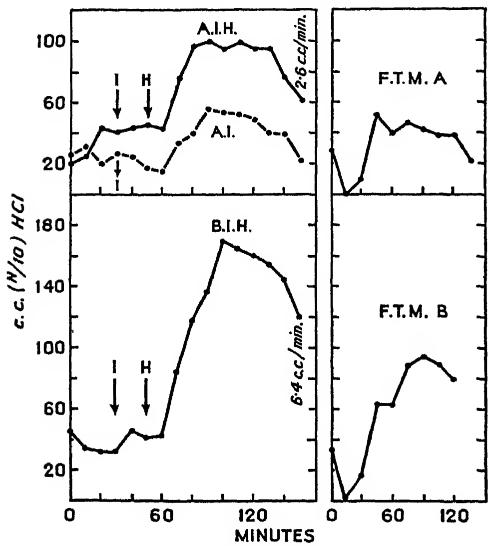


FIG. 517.—Gastric Secretory Response to Insulin plus Histamine. (From an experiment by Dr. Morton Gill.)

Ordinate: c.c. N/10 HCl per 100 c.c. of gastric juice or gastric contents.
UPPER RECORDS (Subject A): Left-hand curves: A.I.H.: response to injection of 7 units of insulin (I) followed 20 minutes later by injection of 0·5 mg. of histamine (H).
A.I.: response to 7 units of insulin (I) only.
Right-hand curves: gruel fractional test meal (F.T.M.) of A.
Note that the response to insulin alone (A.I.) is smaller than the response to insulin+histamine (A.I.H.).
These curves are from subject A with a normal stomach and normal mucous membrane gastroscopically; the volume of secretion averaged 2·6 c.c./min. after insulin+histamine injection.
LOWER RECORDS (Subject B): Left-hand curve: B.I.H. response to injection of 7 units of insulin (I) followed 20 minutes later by 0·5 mg. of histamine (H).
Right-hand curve: fractional test meal (F.T.M.) of B.
Both curves are from subject B in whom the mucous membrane of the stomach appeared gastroscopically thickened and congested. Injection of histamine+insulin produced a higher volume of secretion (6·4 c.c./min.) and a higher peak acidity than in the normal subject (A.I.H.).
The F.T.M. for B also shows a greater response than for A.

Influence of Emotional State on Stomach.[1]—Recent studies on a healthy subject ("Tom") with a chronic gastric fistula have shown that the stomach in man responds very readily to changes in the emotional state. The degree of vascularity of the mucosa was determined by direct examination of its colour or by means of a sensitive thermocouple; alterations in gastric motility were suitably recorded. Sudden fear, for example, produced by the entry of an "irate doctor muttering imprecations" caused sympathetic overaction shown by marked vasoconstriction and a resulting decrease in acid secretion (Fig. 518).

[1] Wolf and Wolff, *Human Gastric Function*, New York 1943.

When the subject was "depressed with sadness and self-reproach" he displayed a decreased secretory and a decreased vasodilator response to the ingestion of 30 c.c. of broth (Fig. 519).

Other kinds of emotional disturbance on the other hand led to parasympathetic overaction. Thus the subject was once unjustly rebuked by his medical employer for slowness and inefficiency in carrying out his job and for overcharging. "Tom" experienced "mounting hostility and resentment"; this emotion was accompanied by progressive engorgement and thickening of the gastric mucosa, increased acid secretion, and increased gastric motility (Fig. 520).

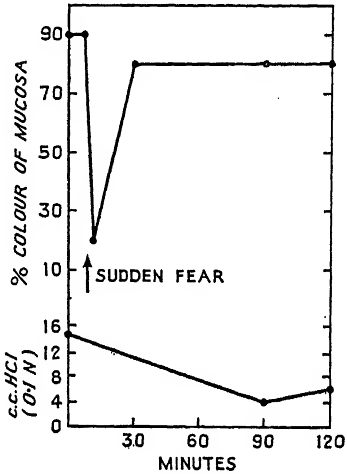


FIG. 518.—Effect of Sudden Fear on Gastric Secretion and Vascularity of the Gastric Mucosa. (Wolf and Wolff, *J. Amer. med. Assoc.*, 1942 *120.*)

Experiment on "Tom" (case of chronic gastric fistula).
Upper record: depth of colour (*i.e.* degree of vascularity) of gastric mucosa as "percentage redness."
Lower record: gastric acidity.
Note vasoconstriction and decreased secretion.

Another striking example may be quoted. Dr. Hoelzel, a student of gastric function, had made some 3000 gastric analyses upon himself before 1928; "the morning gastric aspiration was as regular as washing the hands or face." On 24th January 1928 there was an attempted robbery at his house in *Chicago* (emphasis must be placed on the locality) and his landlady was shot dead. Dr. Hoelzel was responsible for the arrest of the culprits, and for the succeeding 10 days he was in a state of acute anxiety lest he might be shot by the revengeful accomplices of the gangsters. Normally Dr. Hoelzel's fasting morning free gastric acidity varied between 0 and 0·13% (of HCl); on the morning of the shooting episode it was 0·26% and it remained above 0·17% during the period of anxiety. He then moved to what he believed was a safe place and his free gastric acidity fell to within his normal range.

Peptic Ulcer.—The cause of peptic ulcer is unknown but it is probable that the acid and pepsin of the gastric juice are factors in initiating and maintaining the ulcer. The common sites of peptic ulcer are: (i) The first part of the duodenum (*duodenal ulcer*). (ii) In the stomach, usually on the lesser curvature a few inches from the pyloric sphincter (*gastric ulcer*).

(i) Patients with *duodenal ulcer* are of a worrying conscientious disposition showing evidence of overactivity of the gastric vagi; thus the stomach is hypermotile and empties rapidly; a large volume of highly acid juice is secreted which is imperfectly neutralized giving a climbing type of curve in the fractional test meal (Fig. 514, A); an excessive volume of highly acid juice is secreted at night. It is thought that the discharge of the very acid juice into the duodenum produces local damage before there is time for adequate neutralization to take place. A similar mechanism may be responsible for

ulcers in the neighbourhood of the stoma of a gastro-jejunal anastomosis (*gastro-jejunal ulcer*).

(ii) In patients with *gastric ulcer*, the fractional test meal shows variable degrees of acidity and high levels may not be attained. It is probable however that these patients have periods of anxiety or other emotional states which stimulate the vagal mechanism leading to intense congestion and hypersecretion; actual hæmorrhage may take place and it is not unlikely that the potent gastric juice may digest and so further damage any abnormal area. It

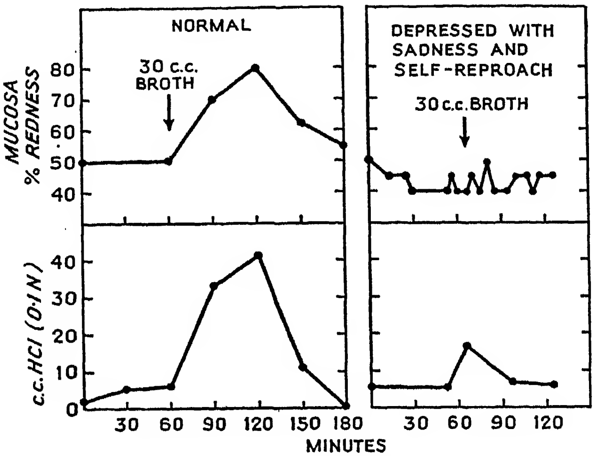


FIG. 519.—Effects of Emotion on Gastric Secretory Response to Food. (Wolf and Wolff, *J. Amer. med. Assoc.*, 1942, *120.*)

Observations on "Tom" (case of chronic gastric fistula).
Records from above downwards: Percentage redness of mucosa; HCl secreted in c.c. of 0·1N HCl per 100 c.c.
Left-hand figure: changes in HCl secreted and vascularity of gastric mucosa (indicated as percentage redness) in response to ingestion of 30 c.c. of broth, with the subject in a normal emotional state.
Right-hand figure: reduced secretory and vascular response when the subject was "depressed with sadness and self-reproach."

is known that hypothalamic lesions, presumably by stimulating the vagi may lead to gastric hæmorrhage, erosion or perforation.

One of the main agents normally protecting the mucosa against the digestive action of the gastric acid and pepsin is the *mucus*. In the patient with the gastric fistula ("Tom") it could be shown that when the mucus was removed from a patch of gastric mucosa and a small injury inflicted, a localized ulcer resulted; such injuries were harmless when the layer of mucus was intact. In predisposed subjects irritant constituents of the diet might perhaps similarly lead to ulceration.[1]

Because vagal overaction may be an important causal factor in peptic ulcer, vagotomy has been carried out on many patients.

RESULTS OF VAGOTOMY.[2]—Both vagi are severed, usually where they

[1] *Pain* of peptic ulcer is discussed on p. 763.

[2] Cameron, *Amer. J. med. Sci.*, 1947, *214*, 202; Moore *et al.*, *J. Amer. med. Assoc.*, 1947, *133*, 741; Walter *et al.*, *ibid.*, 1948, *136*, 742.

lie on the surface of the lower part of the œsophagus. When the resection is completely performed no gastric secretion occurs in response to injection of *insulin*, the sight or taste of food, or emotional disturbances. The chemical phase of gastric secretion and the response to histamine are probably unaffected. In duodenal ulcer cases the volume and acidity of the night secretion are reported to be reduced; if these observations are substantiated they would indicate that reflex vagal secretion can occur even during sleep. Initially gastric tone and motility are greatly diminished; the emptying time of the stomach is increased from 2–3 hours to 12–24 hours or longer. The patient complains of fullness, occasionally of nausea, and belches up large volumes of gas, often foul-smelling. If slight pyloric narrowing is present before the operation the weakened stomach movements resulting from vagotomy may lead to prolonged retention of the gastric contents. The emptying time of the gall-bladder is unaltered. Surprisingly, diarrhœa has been frequently noted (although the vagi are motor to the small intestine and proximal colon). Gastric motility is said to return to normal after about one year.

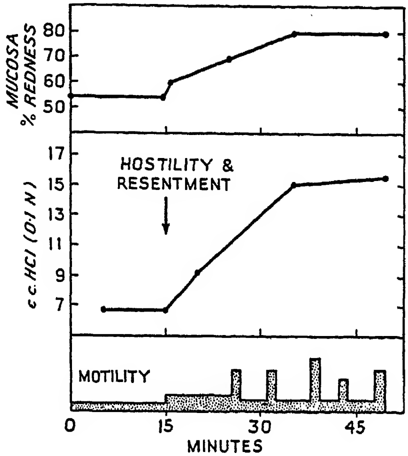


FIG. 520.—Effects of Emotion on Gastric Motility and Blood Supply of Gastric Mucosa. (Wolf and Wolff, *J. Amer. med. Assoc.*, 1942, *120*.)

Observations on "Tom" (case of chronic gastric fistula).
Upper record: percentage redness of mucosa.
Middle record: gastric acidity in c.c. of 0·1N HCl per 100 c.c. of gastric juice.
Lower record: motility of stomach.
In association with the feeling of "hostility and resentment" there is an increase in gastric motility, increase in volume of HCl secreted, and increased vascularity of the gastric mucous membrane.

The results of vagotomy prove that the vagi do *not* carry afferent fibres from the alimentary canal. After the operation the threshold for pain produced by distension (with a balloon) of œsophagus, stomach, duodenum, or small intestine is unchanged. The patient can vomit and experiences appetite, hunger, loss of appetite, fullness of the stomach, and pain due to an unhealed ulcer (p. 763).

The clinical results in suitably selected cases have been encouraging.

SECRETION AND FUNCTIONS OF PANCREATIC JUICE [1]

The pancreas is a dual organ; the externally secreting alveolar tissue forms *pancreatic juice*; the islets of Langerhans (p. 909) form an internal secretion, *insulin*.

COLLECTION OF PANCREATIC JUICE.—In *animals* the secretion may be

[1] Thomas, *External Secretion of the Pancreas*, Springfield, Ill., 1950.

collected in *acute* experiments by introducing a cannula into the pancreatic duct. A *permanent* pancreatic fistula is prepared by transferring the duct papilla with a portion of the surrounding duodenal mucous membrane to the surface of the abdominal wall.

In *man* [1] pancreatic secretion has been studied in cases of pancreatic fistula, following some upper abdominal operation. More indirect information is obtained by aspirating the duodenal contents especially after injecting substances which stimulate pancreatic secretion.[2] As the duodenal fluid is a mixture of pancreatic juice, bile, and succus entericus in varying proportions, the results must be interpreted critically.

Composition of Pancreatic Juice.—Pure pancreatic juice is a colourless, transparent, viscous fluid which is alkaline (pH 8·4) due to the presence of $NaHCO_3$ (0·3–0·65%). In man it is estimated that 500–1200 c.c. containing about 8 g. of NaCl are secreted in 24 hours. The juice contains proteins (of the albumin and globulin types) and three main enzymes, *trypsin*, *lipase*, and *amylase*. The juice which is secreted in response to vagus stimulation or injection of pancreozymin is viscid and rich in enzymes; the juice produced by injection of secretin is more watery, highly alkaline in reaction, and poor in enzymes. Fig. 59 shows the ionic pattern of pancreatic juice.

Functions of Pancreatic Juice.—(i) $NaHCO_3$ is present in such concentration that it neutralizes the acid present in an approximately equal volume of gastric juice. The reaction of the small intestine is thus prevented from becoming excessively acid (cf. p. 862).

(ii) *Trypsin* is a powerful proteolytic enzyme which acts best in an alkaline medium but retains its activity in a slightly acid medium; it breaks down proteins to polypeptides and, given time, may liberate certain amino-acids, *e.g.* leucine, tyrosine; it also curdles *milk*. Juice collected directly from the pancreatic duct (without coming into contact with the duodenal mucosa) is inactive, because it contains not trypsin, but a precursor called *trypsinogen*. The latter is activated by *enterokinase*, a constituent of the succus entericus. (Cf. thromboplastin, p. 142.) Other protein-hydrolysing enzymes, *e.g.* *chymotrypsin* (which mainly resembles trypsin in its action) and "erepsin" (p. 879) may also be present.

(iii) *Amylase* converts all forms of starch rapidly into maltose (p. 835).

(iv) *Lipase* splits neutral fats (triglycerides) to diglycerides, monoglycerides, and free fatty acids, and glycerol. *Bile* greatly increases the activity of the enzyme which may be secreted largely in the form of an inactive precursor (prolipase).

Regulation of Pancreatic Secretion.—1. Nervous Phase.—There is a small *continuous* secretion of juice in the fasting animal; the mechanism is unknown. Within a few minutes of taking food in the intact animal the flow of pancreatic juice is increased for 10–20 minutes. This response depends on reflex stimulation of the *vagi* (from the mouth). Similarly stimulation of the peripheral end of the cut vagus (using the technique described on p. 777) sets up a flow of enzyme-rich juice. This activity is accompanied by shrinkage of the gland cells and disappearance of the contained granules.

2. Chemical Phase.—The introduction of certain substances, particularly

[1] M'Caughan, Senner and Sullivan, *Arch. int. Med.*, 1938, *61*, 739.

[2] A combined stomach and duodenal tube may be used (Rüsch); the gastric contents are aspirated by continuous suction to prevent them passing into the duodenum.

acid, fat, and bile, into the stomach or intestine of experimental animals calls forth a further secretion of pancreatic juice. They all act only on coming into contact with the duodenal mucosa.

(i) *Acid.*—0·5% HCl is very effective ; weaker solutions produce a smaller response. Secretion begins after a latent period of 2–3 minutes (Fig. 521), rapidly reaches its peak after 30 minutes, and goes on at a diminished rate for another hour. This so-called " acid juice " (*i.e.* pancreatic juice secreted in response to acid) is highly alkaline in reaction and has a lower enzyme content than vagus juice.

(ii) *Fat* is as effective as weak acid.

(iii) *Bile.*—The active principle is the bile salt. There is a considerable degree of correlation under natural conditions between the rate of inflow of bile into the intestine and the rate of pancreatic flow.

There is no evidence, however, that either achylia gastrica or exclusion of the bile from the intestine is associated with a decreased secretion of pancreatic juice.

(iv) Water, even in large volumes, and irritants (pepper, mustard) are ineffective. *Alkali* inhibits pancreatic secretion.

Mechanism of Chemical Phase.—The chemical phase is due to the release of one or more hormones (*secretin, pancreozymin*) from the intestinal mucosa into the portal blood ; they reach the pancreas via the systemic circulation ; nervous connections play no part in the reaction. The experimental proof is as follows :

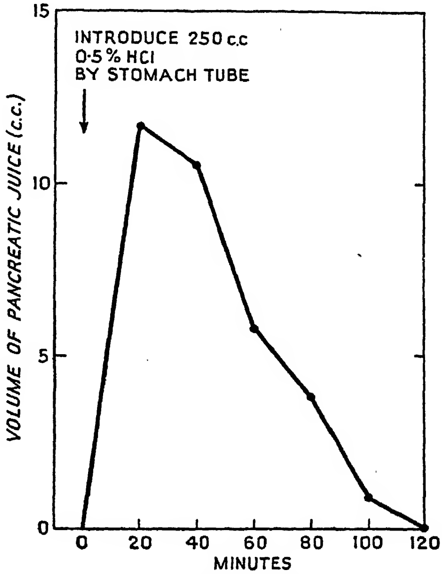


Fig. 521.—Chemical Phase of Pancreatic Secretion. (Data from Pavlov, *Work of Digestive Glands.*)

Dog with pancreatic fistula. The volume of pancreatic juice secreted in the *first hour* after the introduction of other substances into the stomach was as follows :

0·5% HCl	26 c.c.
0·1% HCl	10 c.c.
oil	10–13 c.c.
red pepper	0
mustard	0
water	0

(i) The spinal cord and sympathetic ganglia in the abdomen are destroyed; a loop of jejunum is isolated and its nerve supply is destroyed without interfering with the blood supply. The *introduction of acid into the jejunal loop* causes a flow of pancreatic juice. As all nervous connections between the loop and the pancreas are severed, and as *intravenous* injection of acid has no secretagogue action, the acid can only act by releasing a hormone into the circulation.

(ii) A transplant of a jejunal loop is prepared ; the tail of the pancreas is also transplanted to the surface. New blood vessels grow into these completely *denervated* transplants which function normally. The introduction

of an HCl solution into the isolated jejunum causes the denervated pancreas to secrete.

(iii) Using cross-circulation methods the portal venous blood leaving the small intestine can be shown to contain a secretagogue substance which stimulates pancreatic secretion.

The identity of the structures in the intestinal mucosa which form and release the hormones has not been determined.

Secretin.—Secretin can be extracted from the mucosa of the duodenum (and to a less extent of the upper small intestine). It is a protein of low molecular weight (5000) resembling gastrin in its chemical properties. It produces a flow of watery, highly alkaline juice (Fig. 522) which is almost devoid of enzymes; no degranulation of the pancreatic cells takes place. It markedly increases the secretion of Brunner's glands of the duodenum; the secretion of bile and of succus entericus is also increased. The blood pressure is unaffected by pure preparations. The action of secretin on the pancreas is greatest when it is injected *intravenously*; given subcutaneously it has a more prolonged but slower and feebler action; it is ineffective by mouth.

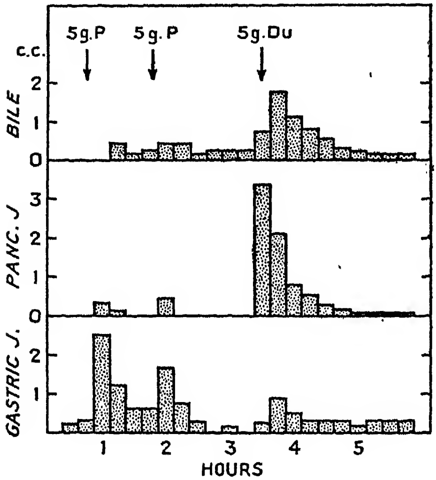


FIG. 522.—Action of Pyloric Extracts (=Gastrin), and Duodenal Extracts (=Secretin) on Secretion of Gastric and Pancreatic Juice and the Flow of Bile. (Komarov, *Rev. Canad. Biol.*, 1943, *1*, 391.). (Cf. Fig. 511.)

Cat. Choralose. Vagi cut.
P.: Pyloric extract (gastrin): it increases the volume and acidity of gastric juice
Du.: Upper duodenal extract (secretin): it increases the flow of pancreatic juice and bile; there is a slight effect on gastric secretion.

The action of secretin in man can be studied by aspirating the duodenal contents; these are influenced by pancreatic activity, but consist of course of an admixture of pancreatic juice, bile, and duodenal juice in varying proportions under different circumstances.[1] Intravenous injection of secretin in man increases up to sixfold the volume of "duodenal fluid" formed (Fig. 523); the bicarbonate concentration increases up to 6 or 7 times so that the total quantity of alkali secreted may be up to 40 times the resting value. The pH of the fluid is about 8·6. The concentration of all the pancreatic enzymes falls, proving that a fluid poor in enzymes is being formed; the total amount of enzyme secreted in a given time is commonly increased (Fig. 524). The interpretation of the results is complicated because it is difficult to assess to what extent the changes recorded are influenced by a simultaneous *increased flow of duodenal juice.*[2]

[1] Ågren and Lagerlöf, *Act. med. scand.*, 1936, *90*, 1. Hammarsten, Ågren and Lagerlöf, *Act. med. scand.*, 1937, *92*, 256. Comfort and Osterberg. *Arch. int. med.* 1940, *66*, 688.

[2] SECRETIN TEST OF PANCREATIC FUNCTION.—On the whole this method of investigation has yielded disappointing results. In cases of *obstruction* of the pancreatic duct

Pancreozymin.[1]—This substance which is also extracted from the upper small intestine produces on injection the same effects as vagal stimulation; *i.e.* there is a flow of juice rich in enzymes, and degranulation of the pancreatic cells occurs. The action of the hormone is not annulled by atropine.

The effects of vagal stimulation and of pancreozymin can be imitated in man by injection of a slowly acting acetylcholine-like substance, *e.g.* mecholyl. Injection of mecholyl (which acts at the vagal terminals) has a slight effect on the volume of duodenal fluid (Fig. 523), but produces a marked and prolonged increase in the output and concentration of all the pancreatic enzymes (by a direct effect on the pancreas) (Fig. 524).

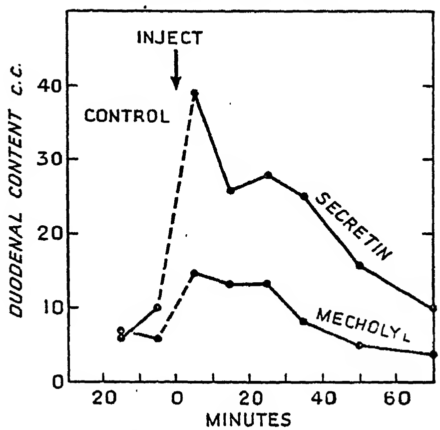


FIG. 523.—Effect of Secretin and of Mecholyl (=Vagal Stimulation) on Volume of Duodenal Contents in Man. (Comfort and Osterberg, *Arch. int. Med.*, 1940, *66*, 693.)

INFLUENCE OF DIET ON PANCREATIC SECRETION.—As the relative activity of the vagal, secretin, and pancreozymin mechanisms probably varies under different circumstances it is not surprising to find that the ingestion of different foods in animals produces a pancreatic juice which varies in volume and composition. *Meat* elicits a large volume of juice rich in alkali and poor in enzyme concentration. Fat or *milk* produces a moderate volume of juice with a low alkali and high enzyme concentration. *Bread* produces an intermediate response (Fig. 525).

CONTROL OF PANCREATIC SECRETION IN MAN.[2]—In a few clinical cases of pancreatic *fistula* attempts were made to study the factors controlling pancreatic secretion in man. The flow was continuous but at a low level between meals; it was markedly increased for about 3 hours by taking food (whether it was rich in protein, fat, or carbohydrate).

Complete Extirpation of Pancreas in Man.[3]—This operation has been successfully performed in cases of malignant disease of the pancreas. The resulting diabetes mellitus is considered on p. 917. The absence of the external secretion produces important digestive disturbances, leading to the loss of large amounts of fat and organic nitrogenous substances. The fæces

---

(*e.g.* by calculus) secretin elicited a poor response, *i.e.* there was only a small increase in the volume of duodenal fluid. In most cases of *chronic pancreatitis* no deviation from the normal range was observed. In *cancer* of the pancreas there was a subnormal secretion of bicarbonate.

[1] Harper and Raper, *J. Physiol.*, 1943, *102*, 115; Harper and Mackay, *ibid.*, 1948, *107*, 89.

[2] McCaughan *et al.*, *Arch. int. Med.*, 1938, *61*, 739; *Arch. Surg.*, 1941, *43*, 269.

[3] Waugh *et al.*, *Proc. Staff Mayo Clin.*, 1946, *21*, 25; Gaston, *New England J. Med.*, 1948, *238*, 345.

are bulky, light in colour, soft, occasionally loose or watery, and sometimes foul smelling; one to three stools are passed daily. The weight of fæcal solids is 80–120 g. per day, or about three times the normal (cf. p. 814).

(i) On a daily fat intake of 100 g., 5–7 g. are normally lost in the fæces; in an operated case on a daily fat intake of 70 and 100 g., 36 and 48 g. were lost in the fæces.

The fat digestion (and absorption) that continues in the absence of pancreatic lipase is due to the action of gastric and intestinal lipase.

The proportion of split to unsplit fat (*i.e.* fatty acid to triglyceride) in the fæces varies, but frequently the *split fat* may constitute up to 80% of

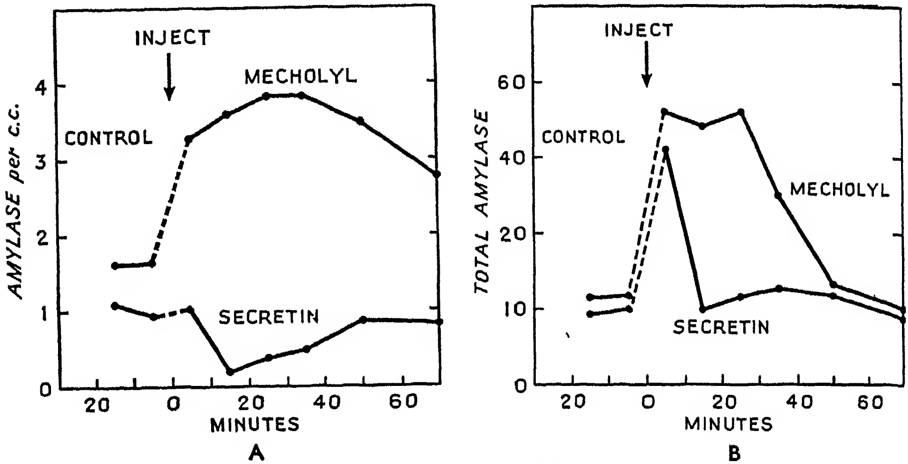


FIG. 524.—Effect of Secretin and of Mecholyl (=Vagal Stimulation) on Concentration and Total Output of Pancreatic Amylase in Man. (Comfort and Osterberg, *Arch. int. Med.*, 1940, *66*, 697.)

Duodenal contents aspirated
Vertical axis in A: *concentration* of amylase per c.c. of duodenal contents; in B: *total* amylase in each 10-minute sample. Amylase in arbitrary units.
Mecholyl produces a more concentrated juice and a larger total secretion of enzyme than secretin. Fall of enzyme concentration (in A) and marked increase of total enzyme (in B) brought about by secretin indicates the secretion of a large volume of fluid.

the total fat lost. This result is surprising in view of the poor absorption that is taking place; it is attributed to the lipolytic action of bacteria in the *colon*; these also continue to act on the fæces *after they have been passed*; of course no fat absorption can take place in the large intestine.

(ii) The normal amount of fæcal nitrogen passed daily is 1·5 g. or less, corresponding to about 10 g. of protein (much of this represents the protein of the fæcal bacteria). In an operated case on a daily protein intake of 100 g., 4–8 g. of N were lost daily, corresponding to 25–50 g. of protein. Pancreatic proteolytic enzymes can then be partially but not wholly replaced by gastric pepsin and intestinal erepsin.

(iii) The digestion and absorption of carbohydrate is normal; thus pancreatic amylase can be wholly replaced by intestinal enzymes.

(iv) The imperfect digestion and absorption of fat and protein may lead

to a loss of 20–35% of the ingested calories. The administration of pancreatic digestive extracts by mouth may halve the loss in the fæces.

(v) In animals pancreatectomy is followed by the development of fatty liver (p. 868).

No hepatic disturbances were noted however in these cases, in spite of the absence of the pancreatic lipotropic substances (p. 868) even in a patient

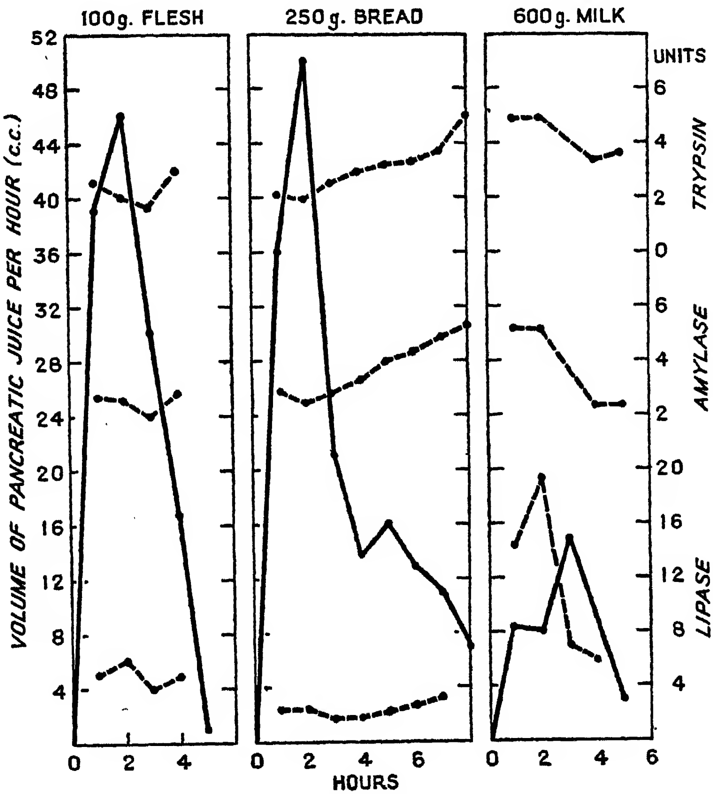


FIG. 525.—Pancreatic Response to different Foodstuffs. (Data from Pavlov, *Work of Digestive Glands.*)

Experiments on dogs with pancreatic fistula. Curves show volume of pancreatic juice (c.c./hour) and concentration of pancreatic enzymes (in arbitrary units) in response to eating meat, bread, and milk.

who refused to take lipotropins. But it is wise to administer extra lipotropins in the form of egg yolk, choline, and methionine; extra calories should be given to make good the fraction lost in the fæces.

Similar observations have been made in children with congenital pancreatic deficiency (fibrocystic disease of the pancreas).[1]

[1] In less severe pancreatic disease the digestive disturbances are correspondingly less marked.

SECRETION AND FUNCTIONS OF SMALL INTESTINE[1]

METHODS.—To investigate the factors controlling the secretions of the small intestine in unanæsthetized *animals*, a loop of intestine is separated and both free ends are sutured into the abdominal wall; the continuity of the bowel is restored; the secretion of the loop can be collected and studied. In *man*, the small intestine can be blocked in two places by means of balloons; by suitable devices the intestinal contents above the proximal balloon or between the two balloons can be aspirated and examined.[2]

DUODENAL JUICE.—*Brunner's glands* are the distinctive histological feature of the duodenum; they secrete an alkaline juice rich in mucus, resembling that formed by the structurally similar pyloric glands. Secretion is increased by *vagus* stimulation, injection of *secretin* (p. 792), or by perfusing the duodenal lumen with a dilute solution of HCl. Transplanted isolated duodenal loops are dry at rest, but secrete briskly after ingestion of food (which of course does not enter the loop); the response to food is partly under nervous and partly under hormonal control.

INTESTINAL JUICE.—*Succus entericus* (like duodenal juice) is an alkaline fluid (Fig. 59) containing mucus and two enzymes, *enterokinase* and *amylase*; certain other important enzymes, which complete the digestion of the food-stuffs, are present in the *lining cells* of the intestine, namely:

(i) *Erepsin*, a mixture of several specific enzymes, acts principally and rapidly on peptones and polypeptides, converting them into amino-acids (it can also break down caseinogen and other proteins slowly); other enzymes complete the breakdown of nucleic acid via nucleotides and nucleosides to liberate purine and pyrimidine bases (p. 879).

(ii) *Invertase* converts cane sugar into glucose and fructose; *maltase* breaks down maltose into two molecules of glucose; *lactase* converts lactose into glucose and galactose (p. 835).

(iii) *Enterokinase* (mainly found in duodenal juice) activates pancreatic trypsinogen (p. 790); certain constituents of intestinal juice increase the activity of pancreatic amylase and lipase.

(iv) *Lipase.*

As with duodenal juice, a flow of intestinal juice follows a meal; it is slight during the first 2 hours, but shows a marked increase in the third hour; it is most marked at the upper end of the gut. The control is mainly chemical and only slightly vagal in origin.

Mechanical stimulation of the intestinal mucous membrane increases the volume and total enzyme output of the small intestine. Colicky intestinal contractions and ingestion of water, glucose, egg, and milk, produce similar results. Local irritants increase the volume of fluid and mucus secreted. Histamine has a very slight effect.

Absorption in the Intestine.—Almost complete absorption of the products of digestion and of other materials (*e.g.* water, salts, vitamins, hæmatinic principle) normally occurs in the small intestine. The mechanism of absorption of the different substances is considered on p. 863 (fat), p. 836 (carbohydrate), p. 879 (protein), p. 151 (vitamin-*K*), p. 997 (calcium), p. 208 (iron).

[1] Florey, Wright, and Jennings, *Physiol. Rev.*, 1941, *21*, 36.
[2] Owles, *Clin. Sci.*, 1937, *3*, 1, 11, 21.

JEJUNO-ILEAL INSUFFICIENCY.—(i) *Excision* of the small intestine must be very extensive before it interferes significantly with digestion or absorption. Thus in one patient almost the whole of the small intestine was resected (for regional ileitis), leaving only the *upper three feet.* Carbohydrate absorption was 99% complete, 70% of the ingested protein was absorbed, but fat absorption showed very wide variations. Thus when 97 g. were ingested only 5 g. were absorbed; when 200 g. were eaten, 80 g. were absorbed; the respective losses in the fæces were thus 92 and 120 g. The lost fat was excreted mainly as fatty acid (cf. p. 794). Calcium absorption was greatly decreased owing to the formation of insoluble calcium soaps; the serum Ca fell with resulting attacks of tetany. The tetany was controlled by cutting down the fat intake and increasing the carbohydrate and giving large amounts of Ca and Vitamin-*D*.

(ii) More severe disturbances result from *gastrocolic fistula,* where much of the food completely short-circuits the small intestine by passing directly from the stomach into the transverse colon. In such patients, in addition to the changes described, there may also be failure of absorption of hæmatinic principle (resulting in pernicious anæmia), and of vitamins (leading to complex multiple vitamin deficiencies).

(iii) Jejuno-ileal insufficiency also occurs in *sprue* and other obscure derangements of the small intestine.

Reduced absorption of foodstuffs from the small intestine may be due to (i) insufficient intake; (ii) inadequate digestion from lack of pancreatic or other juices; (iii) lack of materials necessary to promote absorption, *e.g.* bile salts; (iv) abnormal state of the wall of the small intestine; (v) inadequate length of available small intestine, *e.g.* after extensive resections or fistula formation.

THE BILE[1]

Bile is alkaline and contains the ions found in extracellular fluid (Fig. 59). The principal *organic* constituents of bile are: (i) Bile pigments; (ii) Bile salts; (iii) Cholesterol and lecithin. The origin and significance of each is discussed below.

Bile can be collected from the *common bile duct* or from the *gall-bladder.* The *organic* constituents in gall bladder bile are present in much higher concentration (6–10 times) than in duct bile; *mucin* is also present.

Bile Pigments. (Fig. 104, p. 189).—The bile pigments are *bilirubin* and *biliverdin.* Bilirubin is formed by the scattered macrophages from the hæmoglobin of destroyed red blood cells. It circulates in the blood stream and is excreted by the liver cells from the vascular capillaries into the bile channels. Biliverdin is formed in the bile passages as an oxidation product of bilirubin. The bile pigments serve no digestive purpose. They are partly excreted in the fæces and urine (cf. pp. 188 *et seq.*).

Bile Salts.—The bile salts are probably formed in the liver; they consist of sodium glycocholate and sodium taurocholate. The bile salts, by their hydrotropic action (p. 863), may help to keep the cholesterol of the bile in solution.

[1] Whipple, *Physiol. Rev.*, 1922, 2, 446.

Cholic acid is a steroid which is formed in the body from cholesterol (Fig. 526). *Glycine* (of glycocholic acid) is present in food proteins and is also readily formed in the body. *Taurine*, $CH_2(NH_2).CH_2SO_2.OH$ (of taurocholic acid), is a sulphur-containing substance which is derived from the cystine present in the food or in the body proteins (p. 882). There are about equal amounts of the taurocholate and the glycocholate in human bile. During fasting, the supply of taurine is limited, and this lowers the taurocholate formed to a minimum.

Influence of Diet—During fasting there is a uniform low excretion of bile salts. It is further lowered by ingestion of a carbohydrate-rich diet. It is raised to a maximum by protein-rich diets. On a normal diet there are great variations day by day, depending on the amount of meat and starch which are present in the food.

Rôle in Intestine.—The bile salts increase non-specifically the *digestive action* of all the pancreatic enzymes—amylase, trypsin, and especially lipase. In the various ways which are detailed on p. 863 the bile salts promote the absorption both of neutral fat and fatty acids. When the common bile duct is *obstructed*, 25–75% of the fat intake is lost in the fæces, mainly in the form of fatty acids. The bile salts also promote the absorption of the fat-soluble vitamins -*A* (and carotene), -*D*, and -*K* (cf. p. 151), and possibly of secretin.

Circulation of Bile Salts.—The bile salts are reabsorbed from the intestine, are carried back to the liver, and are then re-excreted in the bile. The evidence is as follows:

(i) No bile salts appear in the fæces, though a trace of a decomposition product of cholic acid (dyslysin) is found.

(ii) If the bile duct is brought up to the surface in a dog and the bile collected, there occurs during fasting a uniform minimal excretion of bile salts, proving that they can be formed constantly in the body. If bile salts are then given by the mouth they are *quantitatively* eliminated from the fistula in 6 hours.

When the bile salts reach the liver they stimulate it to secrete more bile (p. 800), *i.e.* they are *cholagogues*.

As bile salts can be formed in the body and are always returned to the liver from the small intestine for re-excretion in the bile, there must be some method of preventing them from accumulating in excessive amounts. The nature of the regulatory mechanism is unknown.

Cholesterol.—Cholesterol is a steroid; its chemical structure is set out in Fig. 526. The main facts about its metabolism will be summarized here:

Sources.—The chief sources of cholesterol in the food are yolk of eggs, liver, kidney, brain, and fats (cream, butter, and meat fat).

Distribution.—It is widely distributed in most of the cells of the body, either as free cholesterol or combined with fatty acids to form cholesterol esters (cholesterides); it is usually found together with *lecithin*. It is probably a constituent of all cell membranes; it is found in the red cells, and large amounts are present in the ovaries, the adrenal cortex, and the grey matter of the nervous system, especially the brain. Most of the cholesterol derived from cellular disintegration is retained in the body and is not eliminated.

EFFECTS OF INGESTION.—Cholesterol taken in the food is mainly absorbed into the intestinal lymphatics. Some is passed in the fæces after being reduced by intestinal bacteria to *coprosterol.*

The cholesterol of the bile may be an excretion from the blood, for in general the concentration in the bile varies with that in the blood. A lipide-free diet may be associated with a low bile cholesterol level; but excess feeding with cholesterol produces only slight changes in the amount of cholesterol in the bile.

SYNTHESIS.—Cholesterol can be synthesized in the body as proved by the fact that the output may exceed the intake over periods of time by as much as 0·3 g. daily. Liver slices readily synthesize cholesterol at a rate which should be sufficient to supply all the body requirements; tracer experiments prove that most of the 27 C atoms of cholesterol come directly from "acetic acid units" (p. 873).[1] Though the liver is the main site, other tissues (but not brain or nerve) may also synthesize cholesterol normally.

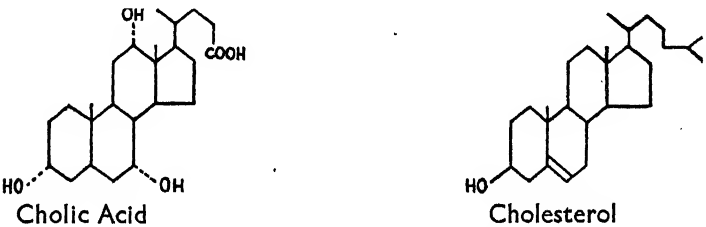


Fig. 526.—Structure of Cholic Acid and Cholesterol.

For full structure of steroid nucleus, see Fig. 662. The projecting lines at C_{10}, C_{13}, represent an attached CH_3 group. Single bonds indicate that the C atom is saturated with H atoms.

BLOOD CHOLESTEROL.—The normal level is 150–250 mg–%. It is equally distributed between the plasma and the corpuscles; one-third is present as free cholesterol, and the remaining two-thirds as cholesterol esters (cholesterides). The ingestion of food rich in cholesterol over a period of time raises the blood cholesterol, but following single meals the changes in the blood in man are variable.

FUNCTIONS.—Cholesterol is the *immediate* precursor of *cholic acid* and of *progesterone*; it is the probable precursor of the *œstrogens* and the *adrenal corticoids.*

PREGNANCY.—Up to the thirtieth week there is a progressive rise in the *free* cholesterol and a corresponding decrease in the *ester* cholesterol to quite low values; the blood returns to normal soon after the puerperium.

IN DISEASE.—The blood cholesterol is said to be low in all forms of chronic *anæmia* and during acute infectious diseases. It is regularly and markedly raised in cases of *thyroid deficiency* (Fig. 623) and decreased in *hyperthyroidism* (pp. 984, 991). Some increase may occur in severe cases of *diabetes mellitus.* In *nephrosis* with extensive œdema, the blood cholesterol is commonly raised, even to 1200 mg–% (p. 114). A patient with *gallstones* who is not jaundiced shows no alteration in the blood cholesterol, so no help in diagnosis

[1] Bloch *et al.*, *J. biol. Chem.*, 1950, *183*, 33; *Recent Prog. Hormone Res.*, 1951, *6*, 111.

can be obtained from such determinations. On the other hand, *jaundice* of the *chronic obstructive type* from any cause is associated with a high blood cholesterol (250–700 mg-%), because the constituents of the bile are passed back into the blood. If the bile is allowed to escape from the body through a fistula, the cholesterol esters in the plasma may almost completely disappear.

Secretion of Bile.—In an animal (or man) with a fistulous opening into the gall-bladder and the common bile duct tied, it can be shown that the secretion of bile by the liver is a continuous process. Bile is secreted at a low pressure; consequently its flow can be readily obstructed. The total volume of bile formed in man is 500–1000 c.c. in 24 hours. Its rate of secretion is modified by the following factors:

(i) About the third hour after the ingestion of food there is an increase in the volume of bile secreted; there is a slight decrease in the flow at night.

(ii) Injection of secretin and stimulation of the vagus nerve have an excitor effect on bile secretion.

(iii) The most powerful *cholagogue* is whole bile or its active principle, the bile acids. If 8–12 g. taurocholic acid are ingested, there is an increase, for 24–48 hours, in the volume of bile which is secreted; the ingested acid can be recovered quantitatively from the bile. Glycocholic acid acts less powerfully.

(iv) Experimental procedures, *e.g.* pulling on the gastro-hepatic omentum or opening and closing the abdomen, reflexly decrease the volume of bile secreted; the administration of an anæsthetic has the same effect.

The results of complete biliary obstruction are fully considered on p. 804.

Functions of the Gall-Bladder.—The gall-bladder is a thin-walled structure; in man its maximum capacity is about 50 c.c. of bile. As mentioned above, the organic constituents in gall-bladder bile are present in 6–10 times the concentration found in the bile in the hepatic ducts owing to the absorption of isotonic saline by the mucous membrane. The valves of Heister [spira valve] at the junction of the gall-bladder and the cystic duct do not resist to any significant extent the flow of bile in or out of the gall-bladder. The terminal part of the common bile duct joins the pancreatic duct in the ampulla of Vater which opens at the apex of the duodenal papilla (Fig. 527); the termination of the common bile duct is controlled by the so-called *sphincter of Oddi*.

Flow and Storage of Bile.—(i) During a *fasting* period the tone of the sphincter of Oddi rises so that it can resist a pressure of 30 cm. of bile. The liver goes on secreting bile steadily; when the pressure in the distended bile ducts rises to about 7 cm., bile begins to flow into the gall-bladder, where it is stored until it is needed during the next meal. Although the gall-bladder can only hold a small volume of *fluid*, it can store the *organic content* of a large volume of liver bile, because 50 c.c. of bladder bile represent the organic content of 300–500 c.c. of liver bile. In this way the gall-bladder acts as a very effective reservoir in spite of its small capacity and prevents the pressure in the bile passages from rising excessively.

(ii) During *periods of digestion* the sphincter of Oddi partially relaxes and now yields at a pressure of about 10 cm. of bile; at the same time the gall-bladder contracts. The pressure in the bile-duct system may rise in consequence to 20 cm.; bile is consequently discharged in a gush into the duodenum.

(iii) If the bile ducts are *obstructed* the pressure in the bile passages in the liver may rise finally to 30 cm., at which point reabsorption of bile into the blood takes place.

Movements of the Gall-Bladder.[1]—These can be studied in animals as follows: The gall-bladder is exposed, emptied of bile, and filled with an iodized oil which is opaque to X-rays; alternatively the clinical method of cholecystography may be employed (p. 802). The following observations can be made radiographically:

(i) *Adrenaline* or *sympathetic* stimulation produces a powerful contraction of the gall-bladder; but stimulation of the vagus, injection of secretin, or the introduction of HCl into the duodenum has no effect.

(ii) Bile salts *increase* the size of the gall-bladder shadow, owing to increased secretion of bile by the liver.

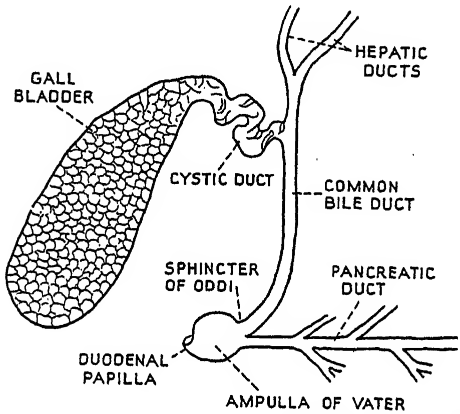


FIG. 527.—Anatomy of Biliary Tract (Barclay, *Digestive Tract*, Cambridge.)

(iii) As each wave of peristalsis in the duodenum is preceded by a wave of relaxation, it has been thought that the alternate relaxation and contraction of the muscle round the termination of the common bile duct might "milk" the bile into the bowel. This is not the case, however, for active peristalsis may take place in the intestine without the gall-bladder being emptied.

(iv) Respiratory movements are unimportant, because all the intra-abdominal structures are uniformly compressed.

(v) *Effect of Food.*—The most effective stimulus to contraction is the presence of large amounts of *fat* in the intestine (particularly *egg yolk*), and to a less extent of *protein*. Expulsion of the oil from the gall-bladder begins in 5–30 minutes, and continues until the organ is emptied after 2–5 hours. The rapidity with which the gall-bladder commences to contract after ingestion of fat is surprising. The same results are obtained after tying the lacteals or division of the vagi and of the sympathetic nerves to the gall-bladder, indicating that some substance *enters the blood stream* from the bowel and, causes contraction of the gall-bladder. The presence of certain foods in the

[1] Ivy. *Physiol. Rev.*. 1934. *14*, 1.

bowel thus not only increases the volume of bile secreted by the liver (p. 800), but also causes partial or complete emptying of the gall-bladder, bile thus being provided in the small intestine to assist the digestive processes.

(vi) *Cholecystokinin.*[1]—Acid extracts of the mucosa of the small intestine contain a substance, *cholecystokinin*, which produces contraction of the gall-bladder (Fig. 528). Ivy claims that cholecystokinin is normally formed and absorbed into the blood stream during intestinal digestion and acts on the gall-bladder.

(vii) Except in response to the presence of large amounts of fat in the bowel, the gall-bladder does not usually empty itself completely, but fills slowly and empties irregularly; this is shown by the fact that on a normal mixed diet, dyestuffs introduced into the gall-bladder are still found there after 1–3 days, but not after 7 days.

Cholecystography.—The biliary system in man can be examined radiographically by Graham's method. When tetraiodo- or tetrabrom-phenolphthalein is injected intravenously or ingested it is excreted by the liver

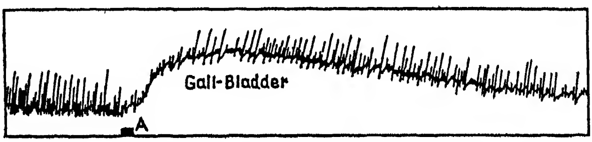


FIG. 528.—Action of Cholecystokinin on Gall-Bladder Contraction. (Ivy, *Medicine*, 1932.)

Record of gall-bladder movements. At point A on base-line an intravenous injection of cholecystokinin was made. The gall-bladder immediately contracted. There was no change in the volume of the liver or in the blood pressure.

from the blood; it passes into the gall-bladder, where the dye is sufficiently concentrated by the absorption of saline to become opaque to X-rays. The size, shape, and behaviour of the gall-bladder can be studied radiographically. The dye which is present in the liver and the bile ducts is too dilute to cast a shadow. It is safer to administer the dye by the mouth in formolized gelatin capsules; when given intravenously, toxic symptoms may be produced; given subcutaneously, local necrosis results. Using this technique it can be shown that *fat* and, to a less extent, *protein* taken by mouth cause rapid emptying of the gall-bladder in 3–8 hours. (Figs. 529 and 530 show the effect of ingesting egg yolk; cf. also p. 801.)

Cholecystography is employed in the diagnosis of gall-bladder disease. If the dyestuff is absorbed from the bowel and yet *no* shadow of the gall-bladder is obtained, we can conclude that the cystic duct is obstructed or the gall-bladder is sclerosed, or that the gall-bladder wall because of disease cannot concentrate the contained bile in the normal way.

Effect of Magnesium Sulphate. Drainage of Biliary Passages.—It is claimed by Meltzer that certain salts cause relaxation of the sphincter of Oddi and tonic contraction of the gall-bladder. Lyon has made use of this observation: a tube is passed into the duodenum, a 33% solution of *magnesium sulphate* is introduced, and the bile is collected at intervals. The first samples are light yellow, and come from the common bile duct; then follows darker and more viscid bile from the gall-bladder; and, finally, light thin bile of

[1] Ivy, *Medicine*, 1932, *11*, 345.

low specific gravity, and containing only traces of cholesterol, is obtained from the hepatic ducts. This technique can be employed to drain the biliary passage or to collect bile for bacteriological and microscopical investigation.

It must be pointed out, however, that magnesium sulphate given by mouth is probably equally effective ; that, studied radiologically, the Lyon technique does *not* cause powerful contraction of the gall-bladder ; and that it seems more reasonable to give fat, which is known to empty the gall-bladder. *Warm olive oil* introduced into the duodenum is a more reliable stimulus to gall-bladder contraction.

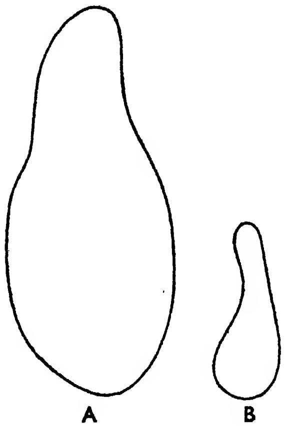


FIG. 529.—Outlines of Human Cholecystograms, about one-half life-size. (Boyden, *Anat. Rec.*, 1926.)

A: gall-bladder 17 hours after ingestion of tetra-iodo-phenolphthalein (vol. = 4 cu. ins.). B: the same gall-bladder 1¾ hours after a meal of 4 egg yolks and ½ pint of cream (vol. = 0·22 cu. ins.). At 180 minutes the gall-bladder was completely emptied, and could no longer be seen radiographically.

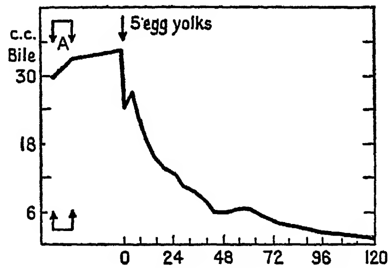


FIG. 530.—Effect on Volume of Human Gall-Bladder of a Meal of Five Egg Yolks mixed with Water. (Boyden, *Anat. Rec.*, 1928, *40*.)

Ordinate: volume of gall-bladder. Abscissa: time in minutes. During the period (10 minutes) indicated by the bracketed arrows, A, the subject smelt a dish of hot fried bacon. The volume of the gall-bladder increased. At 0, he ingested 5 egg yolks. Note in this case that the volume of the gall-bladder decreases markedly in 2 minutes. The eggs begin to enter the duodenum within 1 minute of swallowing.

Removal of the Gall-Bladder.—This operation gives rise in man to disadvantageous consequences :

(i) The bile ducts become dilated to accommodate to some extent the bile which is continuously secreted by the liver.

(ii) If the tone of the sphincter of Oddi is high, the pressure in the biliary passages rises until it may equal or exceed the secretory pressure of the liver cells and thus interfere with their activity.

(iii) If the tone of the sphincter is low (as it often is for a time after cholecystectomy), bile dribbles into the intestine when it is not needed and is consequently wasted.

The importance of the reservoir action of the gall-bladder is illustrated by

the following experiment. If the common bile duct is tied, and the gall-bladder removed, jaundice appears in 3–6 hours; but if the gall-bladder is left it can store so much bile pigment newly secreted by the liver that (after tying the bile duct) jaundice does not develop for 36–48 hours. The rise in retained plasma bilirubin is similarly more marked and more rapid in the animal deprived of its gall-bladder (Fig. 531).

Results of Complete Biliary Obstruction.—Complete obstruction of the bile ducts produces results which are due to:

(1) *Absence of bile from the bowel.* This leads to impaired digestion and reduced absorption of fats and corresponding changes in the fæces (p. 798); reduced absorption of vitamin-*K* leading to a fall of plasma pro-

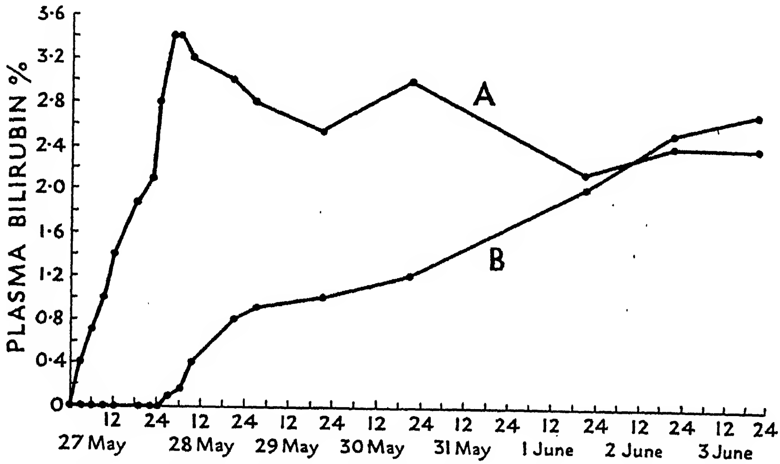


FIG. 531.—Storage Capacity of Gall-Bladder. Effects on Plasma Bilirubin Accumulation of Obstruction of Common Bile Duct in Dogs. (Redrawn from Mann and Bollmann, *J. lab. clin. Med.*, 1925, *10*, 681.)

Ordinate: plasma bilirubin in Van den Bergh units. At 0 in both animals tie common bile duct. In the experiment illustrated by the upper curve (A) the gall-bladder was also removed. Note the more rapid rise of plasma bilirubin in A than in B, owing to loss of storage function of gall-bladder.

thrombin and hæmorrhages (p. 151); defective hæmoglobin formation and anæmia of obscure origin; reduced absorption of fat-soluble vitamins-*A* (and carotene) and -*D*.

(2) The effects on the body cells of *retention of bile in the blood* and tissue fluids.

The *retention of bile* leads to jaundice (p. 191), loss of appetite, and injury to the gastric mucosa; peptic ulcers (usually duodenal) develop frequently, and may be responsible for death from hæmorrhage or perforation. The heart rate is slowed. The blood level of all the organic bile constituents (bile pigments, bile salts, and cholesterol) is increased; bile pigments and bile salts appear in the urine.

(3) The *injury to the liver* resulting from biliary obstruction.

The liver varies in size, is stained with bile pigment and is smooth and

firm. The bile canaliculi are dilated, the hepatic cells are atrophied (mainly round the portal canals) and there is connective tissue overgrowth. The impairment of liver functions leads to the signs of *parenchymal failure* which are detailed on p. 829.

The composition of the *diet* markedly influences the clinical state. In animals on a carbohydrate-rich diet, life may be prolonged for a year or more; an exclusively meat diet may, however, prove fatal within one week. It is pointed out on p. 828, that a high glycogen content protects the liver cells against the harmful effects of various toxic agents.

Bile Fistula.—A complete bile fistula (in which all the bile is passed through an artificial opening to the exterior) results in *loss of bile from the body* with the results summarized on p. 804, (1); there is progressive impairment of bile secretion by the liver. There is no compression of the liver cells and therefore *no parenchymal failure.* The usual results of *loss of water and electrolytes* occur (if these are not replaced).

MECHANICS OF ALIMENTARY CANAL[1]

The movements of the alimentary canal in man are best studied by means of X-rays after the ingestion of a meal which contains opaque barium salts.

Swallowing.—This occurs in three stages; the first stage is *voluntary,* the other two are *reflexly* produced.

(1) After mastication, the food is rolled into a bolus, which lies in the curve of the tongue. Swallowing commences by voluntary contraction of the mylohyoid muscles, which throw the bolus back between the pillars of the fauces on to the post-pharyngeal wall. This region of the pharynx has a rich sensory innervation from the glossopharyngeal nerves; when the local nerve endings (and also those in the soft palate and epiglottis) are stimulated, afferent impulses are set up which reflexly (via the so-called *deglutition centre* in the medulla) produce the complex coordinated movements occurring in the involuntary phases of swallowing.

(2) The soft palate is elevated and thrown against the post-pharyngeal wall to close off the nasal cavity. The larynx rises with the elevation of the hyoid, and the pharynx is practically obliterated. The vocal cords are approximated, and breathing is momentarily inhibited. The posterior pillars of the fauces approximate to shut off the mouth cavity. The pharynx reopens to permit the passage of the bolus; the epiglottis guards the laryngeal opening, until the bolus reaches the œsophagus which simultaneously opens up to receive it. Aspiration of the food into the larynx is also prevented by the associated reflex apnœa.

(3) The bolus is then propelled along the œsophagus by peristaltic waves in its muscle coat. Gravity plays little part in this process, as the rate of progress along the œsophagus is not affected by posture; it is as rapid in the supine as in the erect position.

The swallowing reflex is temporarily abolished by anæsthetising the pharynx with cocaine; it is deranged by lesions of the medulla oblongata

[1] Alvarez, *Mechanism of the Digestive Tract,* 2nd edn., New York, 1928; *Introduction to Gastroenterology,* 1939. Barclay, *The Digestive Tract (Radiological Study),* 2nd edn., Cambridge, 1936.

or of the ninth and tenth nerves. Food may then regurgitate into the nose or be aspirated into the larynx.

Cardiac Sphincter.—The last inch or more of the œsophagus, including the whole of the abdominal portion, is sphincteric in its action. The muscle coat is here thicker, and prominent longitudinal rugæ of the mucous membrane are visible. When swallowing is not occurring, the sphincter is usually in a state of tone and its walls are tightly in apposition; sometimes it is patulous (relaxed) at rest.

Within 5 or 6 seconds of swallowing, the bolus reaches the lower end of the œsophagus, where its progress may be momentarily arrested by the resistance of the sphincter if it is closed. The sphincter then relaxes, but its lumen may still remain considerably narrower than that of the rest of the œsophagus, so that food trickles slowly into the stomach.

When the whole of the meal has been eaten, the cardia usually closes, thus preventing the regurgitation of food, gastric juice, and swallowed air, which would otherwise occur because of the positive intragastric pressure. Increased *acidity* of the gastric contents is alleged to cause closure of the cardia; the same effect is produced by a rise of intragastric pressure.

INNERVATION OF ŒSOPHAGUS.[1]—The *upper* third of the œsophagus is lined by a *striated* muscle coat, the *lower* two-thirds by *non-striated* muscle. The vagi supply the whole œsophagus including the cardiac sphincter; the sympathetic probably acts on the sphincter only. The effects of section of the vagi on œsophageal movements have been carefully studied in the dog. If the right vagus is cut below the origin of the recurrent laryngeal (which is alleged to supply the striated upper third of the œsophagus) and the left vagus is cut in the neck, the lower two-thirds of the œsophagus become *dilated* and *peristalsis ceases*; the *tone of the cardiac sphincter is generally reduced.* The food tends to accumulate in the paralysed region; if the animal is placed in the erect position the raised hydrostatic pressure may force some food into the stomach. Sometimes a further swallow bulges the œsophagus, some food enters the stomach and may be pushed back again so that some of the contents impinge on the pharynx setting up reflex vomiting. From time to time the seemingly paralysed œsophagus *contracts as a whole,* vomiting resulting. When vagotomy is performed below the level of the hilum of the lungs the lower quarter of the œsophagus becomes paralysed.

Section of the sympathetic in the dog does not significantly affect the activity of the œsophagus; sphincter tone is unaffected.

These observations prove that peristalsis along the œsophagus depends on the integrity of the *extrinsic* nerves, the vagi (preganglionic fibres). The excitor neurones (corresponding to Auerbach's plexus) cannot mediate peristalsis. The denervated œsophagus can however contract, presumably in response to local distension. The sympathetic has no effect on œsophageal movements. In the dog the vagi contain both motor and inhibitory fibres to the cardiac sphincter, the former predominating; the sympathetic control is unimportant. The state of the œsophagus after vagotomy in the dog closely resembles that found in human cardiospasm (p. 815).

Reverse peristalsis may occur in the œsophagus, and is responsible in part for belching and acid regurgitation into the pharynx.

If the intragastric pressure is excessively raised by rapid air-swallowing

[1] Mann *et al., Amer. J. Physiol.,* 1947, *149,* 429.

(*aerophagy*) or by the evolution of CO_2 from ingested $NaHCO_3$, the resistance of the cardiac sphincter may be overcome, and gas is expelled by the mouth.

Nature of Peristalsis.—Bayliss and Starling defined *true* peristalsis as a coordinated reaction in which a *wave of contraction preceded by a wave of relaxation* passed down a hollow viscus; the contents of the viscus as they are propelled along would thus always enter a segment which had actively relaxed and enlarged to receive them. This type of movement was thought to be responsible for transferring the contents of the alimentary canal from the œsophagus through the stomach, small and large intestine, and finally to the anus. Later work, especially by Alvarez, makes it improbable that a wave of relaxation regularly precedes the wave of contraction. It is necessary also to explain why the wave of contraction travels normally caudally rather than towards the mouth—in other words, why there is a sort of " law of forward conduction" in the bowel. No satisfactory reason has been discovered, though much is written about a normal forward " gradient " in the bowel, *i.e.* some inherent condition of the bowel neuromusculature which " directs " the peristaltic waves in the normal direction. If a portion of small intestine is excised, the two ends reversed and the continuity of the bowel restored by sutures, it acts as an obstruction because peristaltic waves cannot be transmitted along it. Fluids may still be drawn through, but solids cannot pass, with the result that a block occurs at the proximal junction.

The passage of peristaltic waves along the œsophagus depends on the continuity of the preganglionic vagal nerve supply but not on the integrity of the muscle coat (p. 806). It is claimed that if the œsophageal wall is divided and the superficial nerve plexus is left intact, the peristaltic wave can still pass normally over the œsophagus.

In the stomach and intestines, however, peristalsis can occur in the absence of extrinsic nervous influences (preganglionic vagus or pre- and postganglionic sympathetic) though it is modified by the activity of these nerves. Thus gastric tone and motility are initially decreased by vagotomy in man (p. 789). After section of the vagi and destruction of the abdominal ganglia of the sympathetic in animals, intestinal peristalsis continues normally for months. Normal gastric and intestinal peristalsis is attributed to a series of *coordinated local nervous reflexes* (involving possibly Auerbach's plexus) in response to the chemical and mechanical stimulation set up by the food.

Isolated strips of intestine suspended in a bath of warm oxygenated Ringer's solution carry out rhythmic movements (though these are probably not identical with true peristalsis); they still occur in such isolated strips after they have been deprived of their intrinsic nerve plexuses.

Movements of the Stomach.[1]—ARRANGEMENT OF MUSCULATURE.—The part of the stomach to the left of the incisura angularis is the *body*, and that to the right the *pyloric* part; the part of the body above the level of the cardiac orifice is the *fundus*. The pyloric part is divided into the *pyloric antrum* or vestibule, and the *pyloric canal*. Functionally the first part of the duodenum (duodenal *bulb* or cap) is associated with the pyloric part.

The stomach has an outer *longitudinal* and an inner *circular* coat; between the mucous membrane and the circular coat is an additional incomplete but well-developed muscular layer which runs from the œsophagus down either side of the lesser curvature and then spreads out in a fan-like

[1] Hurst, *Medical Essays and Addresses*, London, 1924; *Brit. med. J.*, 1925, i, 145

Innervation.—Stimulation of the *vagi* produces variable effects on the stomach; contraction is obtained when the initial tone is low, and relaxation if the initial tone is high; sometimes mixed effects may be seen, *e.g.* contraction of one part of the stomach and relaxation of another. In man the vagi exert a continuous stimulating influence on gastric tone and movements as shown by the effects of vagotomy (p. 789).

Gastric Tone.—The stomach displays postural activity or *tone*. When it is empty its walls are firmly in apposition. When food enters the stomach, the muscle fibres are elongated to enlarge the size of the cavity uniformly in order to accommodate the new contents, without much change of internal pressure (cf. p. 768). With an opaque meal of average composition, the stomach assumes varying shapes. Most commonly it is J-shaped (Fig. 532); the body forms a vertical tube and the pyloric part forms a horizontal or slightly ascending segment which is turned to the right. Sometimes the stomach is shaped like a *steerhorn*; the gastric axis is more oblique, and the pylorus is almost the lowest point on the lesser curvature (Fig. 533). Fig. 534 shows a small stomach, often inaccurately called a "hypertonic" stomach. Fig. 535 illustrates a long stomach, often inaccurately called a "dropped" stomach. Both these two extremes are within the range of the normal. Variations in gastric tone are easily produced reflexly, leading to considerable changes in the position of the lower border of the stomach, *e.g.* by slamming a door during the X-ray examination. The normal tonic activity of the stomach wall not only grips its content but also opposes the influence of gravity; as filling occurs the stomach expands almost wholly *laterally*.

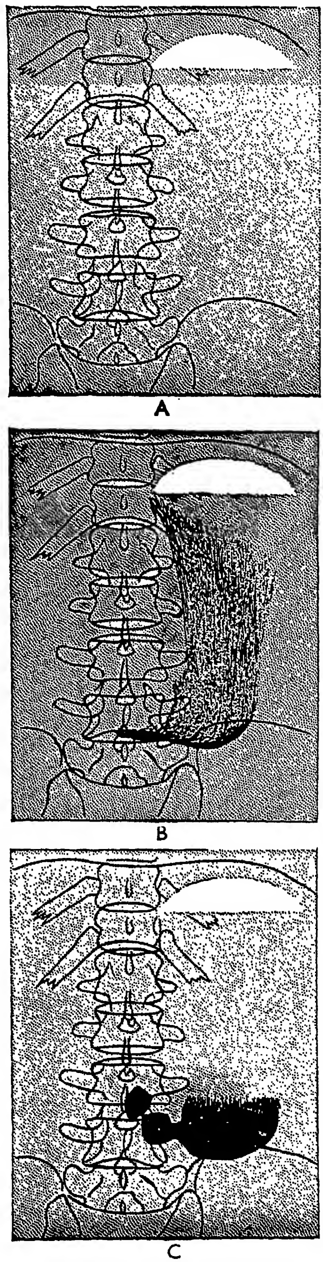


Fig. 536.—Passage of Fluid Meal into a Full Stomach. (Barclay, *Digestive Tract*, Cambridge.)

A. Normal stomach which contains ordinary non-opaque food; the crescentic air-space at the top alone is seen.
B. A watery suspension of barium is given. This flows through and around the non-opaque food, giving a veil-like appearance.
C. The opaque food settles out and lies at the bottom of the stomach.

Foods pass through the stomach to the pyloric antrum roughly in the order in which they are swallowed; but if a heavier food succeeds a lighter, it sinks in the stomach contents till it finds its own level. This fact is well illustrated in Fig. 536. A watery suspension of barium is seen entering a stomach containing ordinary food. The heavy material sinks gradually and finally settles as an opaque mass in the most dependent part of the stomach (Fig. 536, C).

GASTRIC MOVEMENTS.—*Peristaltic waves* begin high up on the body of the stomach and follow one another about three times a minute; they pass rapidly towards the pylorus which at this stage is tonically contracted. The waves are normally gentle and not very deep; they resemble "the waves of a quiet sea rather than breakers beating on the rocks." There is no general vigorous churning-up of the food with the gastric secretions; the surface of the stomach is merely moulded. The peristaltic indentations become deeper as they approach the pylorus (see Fig. 532), but do not pass beyond it into the duodenum.

The behaviour of the *pyloric sphincter* is variable. Five to fifteen minutes after the taking of food it begins to relax occasionally to permit a small volume of chyme to escape; generally such relaxation occurs when a wave of contraction involves the pyloric antrum. When the stomach has emptied to some extent, the sphincter may also relax between the peristaltic waves. Food may pass through the pylorus in the absence of gastric peristaltic waves, presumably when the steady pressure in the stomach exceeds that in the duodenum. Conversely, if the intraduodenal pressure exceeds that in the stomach, or if anti-peristaltic waves occur in the duodenum, intestinal contents may regurgitate into the stomach (cf. p. 782).

The mechanism controlling the pylorus is unknown. According to Barclay, if the small intestine is emptying rapidly, the pylorus opens more readily; on the other hand, if the small intestine becomes overloaded, the pylorus closes and gastric peristalsis is inhibited. The presence of acid in the stomach or duodenum does *not* regulate the activity of the pyloric sphincter. In *achylia gastrica,* in which condition no free HCl is present in the stomach, the organ empties itself, as a rule, with excessive rapidity, and the pyloric sphincter tends to be patulous; but a hypermotile rapidly emptying stomach may also be one which is secreting large amounts of acid (p. 787).

With a meal of gruel the stomach is usually empty in about 2½ hours (Fig. 514). With the heavier and more solid barium meal, the stomach should be completely empty within 6 hours at the longest. *Fat* in the food (probably by releasing enterogastrone) has an inhibitory effect on gastric motility (p. 780). The influence of the emotional state on gastric motility is considered on p. 786.

The *fasting* stomach shows rhythmic variations in tone, which increase and then diminish the pressure within the organ; these occur at the rate of three a minute. At intervals a series of powerful contractions occur which last about 30 seconds, and are associated with the appearance of the sensation of *hunger* (*hunger pains*).[1] These changes can be demonstrated by introducing into the stomach, via the œsophagus, a rubber balloon coated on the inside with barium paste; both graphic and X-ray records of the stomach movements can thus be obtained. It is uncertain whether the sensation of hunger

[1] Carlson, *Hunger in Health and Disease*, Chicago, 1916. Barclay, *Lancet*, 1922, ii, 261.

is caused by these strong gastric contractions. A distinction should also be drawn between hunger and appetite.

Vomiting.[1]—The phenomena of vomiting are as follows: Nausea is first experienced, the secretion of saliva is increased, and the breathing becomes deep, rapid, and irregular. *Retching* may occur, which consists of simultaneous incoordinated spasmodic contractions of the respiratory muscles; the diaphragm, for example, descends when the expiratory muscles contract. The glottis closes and remains shut till the expulsion of the vomited material is effected. The pyloric part contracts firmly, and at the same time the body of the stomach relaxes so that the gastric contents are forced into it; anti-peristalsis may sometimes take place in the stomach, but it is unimportant. The flaccid stomach is compressed by the raised intra-abdominal pressure resulting from the simultaneous descent of the diaphragm and the contraction of the abdominal wall. The cardiac sphincter is inhibited and the gastric contents are therefore driven into the dilated œsophagus. Some of this material is at once expelled from the mouth; some is moved up and down the œsophagus. Towards the end of the act of vomiting, the diaphragm relaxes, *i.e.* ascends, and all the *expiratory* muscles and the abdominal wall contract. As the glottis is closed the intrapulmonary pressure becomes *positive*. The œsophagus is thus compressed; it may also actively contract throughout its length or a wave of anti-peristalsis may pass over it; its contents are thus emptied into the mouth. The palate is raised to shut off the nasal cavity from the throat.

The complex series of movements which occur during vomiting are controlled by a group of nerve cells, in close relation to the vagal nucleus, called the *vomiting centre*. In the dog it is situated just cranial to the calamus scriptorius. It is close to, but distinct from, the respiratory centre. If this centre is damaged, the injection of apomorphine (which is a "central" emetic) no longer induces vomiting. Afferent impulses to produce vomiting may arise in the stomach and other parts of the gastro-intestinal tract, vestibular apparatus, heart, and other organs. Certain drugs and poisons make the centre more sensitive, so that the normal impulses which reach it are sufficient to cause vomiting. In partial asphyxia, and when the intracranial pressure is raised, the vomiting centre is stimulated.

Movements of the Small Intestine.—In the first part of the duodenum the food forms a definite shadow, filling it out into a triangular cocked-hat shape, known to radiologists as the *duodenal cap* or *bulb* (Fig. 532). The shadow remains in this position while food is present in the stomach; no peristalsis occurs in this segment and its shape is often unchanged for long periods. It may empty occasionally by means of a general contraction which, if the pylorus is closed, pushes the food into the second part of the duodenum. Beyond the cap very active intestinal movements take place. It is difficult to see radiographically in man what is going on, but beyond the sharp, well-defined shadow of the first part of the duodenum there is an ill-defined and very diffuse outline indicating that the meal has been rapidly broken up and hurried round the duodenal loop into the jejunum. This fragmentation is perhaps brought about by active contraction of the muscle fibres in the mucous membrane. Anti-peristalsis is occasionally observed in the duodenum. Beyond the duodenojejunal flexure only a cloudy opacity

[1] Hatcher, *Physiol. Rev.*, 1924, *4*, 479.

is seen till the meal arrives at the ileo-cæcal sphincter about *three and a half hours* after it began to leave the stomach.

In *animal* experiments two kinds of intestinal movements can be recognized: *peristalsis* and *segmentation*.[1]

PERISTALSIS.—The chyme is passed along the small intestine by a series of frequent and very rapid peristaltic waves which carry it on a few inches at a time. Each wave lasts about a second, and is followed by a quiescent period of a few seconds to a few minutes.

SEGMENTATION.—A length of small intestine becomes divided by a series of constrictions into a number of small segments which are ballooned at the centre. Each segment is then divided in its middle (*i.e.* at the point of maximum distension), and the previous constrictions disappear. This process occurs about six times a minute. Segmentation causes no forward movement of the intestinal contents; it helps to mix the chyme with the intestinal juices. and so aids digestion; by bringing the chyme into intimate contact with the villi it aids absorption. It helps to squeeze blood and lymph out of the bowel, and so aids their return to the thorax. [For *movements of villi*, see p. 865.]

The vagi end in Auerbach's plexus, which lies between the circular and longitudinal muscle coats (Fig. 447). Stimulation of the vagi increases the tone of the intestine and renders the movements more active (after an initial temporary inhibition). Stimulation of the sympathetic diminishes tone and arrests the peristaltic movements.

The intestinal movements may be influenced by the *ductless glands*; thus diarrhœa is a common symptom of exophthalmic goitre, and constipation of myxœdema.

Ileo-Cæcal Sphincter [Valvula coli].—This consists of the fusion of the circular and longitudinal fibres at the end of the ileum and in the adjoining cæcum; in addition, some independent specialized circular fibres are developed at the apex. On inspection in cases of cæcal fistula the sphincter appears as a smooth scarlet elliptical or hemispherical papilla about 1·8 cm. in diameter, which projects 1 cm. above the pink folded mucosa of the cæcum. The orifice is normally closed, and forms a dimple in the centre of the papilla. Each time a peristaltic wave passes over the terminal part of the ileum, the sphincter relaxes, about 2 c.c. of fluid fæces mixed with a little gas escape, and then the sphincter closes again.[2]

In man, the main function of the sphincter is to prevent the contents of the ileum from passing too rapidly into the cæcum; thus an opaque meal may be visible in the terminal ileum for about one hour before any appreciable quantity enters the cæcum. This delay promotes more complete digestion and absorption of the intestinal contents; as a result the chyme which enters the cæcum contains very little nutrient material.

Peristaltic waves over the terminal ileum are generally infrequent, but when food enters the stomach a *gastro-ileal reflex* is set up, which produces active peristalsis at the extreme end of the ileum; the chyme rapidly passes

[1] If the abdomen of an animal is opened in a bath of warm saline, swaying or *pendulum* movements of the intestine can also be seen.

[2] In animals, section of the sympathetic causes permanent relaxation of the sphincter, and the contents of large and small intestine intermingle freely. Stimulation of the vagus is said to have no effect on the sphincter.

into the cæcum. In man, anti-peristalsis in the ascending colon is feeble, and so there is little tendency for fæces to regurgitate into the ileum. The sphincter, however, is not capable of resisting considerable pressures. When a barium enema is administered at a pressure of 18 inches H_2O, or when powerful retrograde movements occur in the large bowel because of some mechanical obstruction, regurgitation into the ileum takes place.

Movements of Large Intestine.—As stated, the chyme reaches the terminal ileum 3½ hours after a meal, and is held up there for about 1 hour. The barium shadow can be first seen in the various parts of the large intestine the following number of hours (on an average) after a meal: Cæcum, after 4½ hours; hepatic flexure, 6 hours; splenic flexure, 9 hours; descending colon, 11 hours; iliac colon, 12 hours; pelvic colon, 18 hours (Fig. 537).

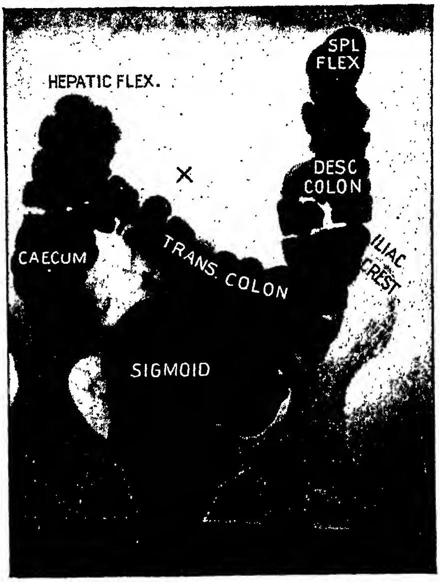


FIG. 537—Normal Radiograph of Large Intestine Particularly Well Filled from End to End. (Barclay, *Digestive Tract*, Cambridge)
× Indicates the position of the umbilicus.

The cæcum and ascending colon are almost entirely *passive*, and fill solely as a result of activity in the ileum, which occurs mainly during and immediately after meals. As a rule, the contents of these parts of the large bowel appear immobile, but the lumen is indented by well-marked folds (*haustrations*) of the intestinal wall; weak peristaltic movements may occur. In animals retrograde waves (*anti-peristalsis*) occur periodically which force the contents of the proximal colon towards the cæcum; such movements have occasionally been detected in man arising near the hepatic flexure, but they are both rare and feeble.

Rhythmic variations of *tone* occur from time to time over all parts of the large intestine. They do not propel the contents onwards, but serve to mix them and aid absorption of water.

The colonic movements are more sluggish by night than by day.

Mass Peristalsis.—After each meal a *gastro-colic* reflex is set up. A brief, powerful peristalsis occurs in the colon, which constricts down, and its contents are driven vigorously onwards. By means of large forward movements of this type, the fæces reach the splenic flexure, then the descending colon, and finally the pelvic colon, where they accumulate. The fæces do not normally pass beyond the pelvi-rectal flexure—*i.e.* the point where the movable pelvic colon joins the fixed rectum at an acute angle. It is uncertain whether a definite sphincter exists here, but local thickening of the

circular muscle fibres is described, and narrowing of the lumen is present at this point in 80% of bodies. The normal rectum is empty, except immediately before defæcation.

About 350 g. of fluid chyme passes through the ileo-cæcal sphincter daily (as observed in cases of cæcal fistula). The average weight of the moist fæces is 135 g. Most of the absorption of water occurs in the cæcum and ascending colon, the contents of which are quite soft; little absorption occurs in the transverse or descending colon. The normal consistency of the fæces is due to further absorption of water which occurs in the pelvic colon.

Defaecation.—The act of defæcation is very much a matter of habit or conditioning as regards time of day, relation to meals, and the frequency of occurrence. Some people defæcate before breakfast, others after; some on getting to work because they had to leave home hurriedly; some in the evening when they can attend to their needs at leisure; some several times a day, others once in several days. A person may refrain from defæcating because he feels shy in a strange house or because the act is painful owing to the presence of, *e.g.* an anal fissure. And there are the people who are anxious to defæcate whenever possible because they think it is good for them. Whenever the time, the place, and the appropriate stimulus are all together, a wave of peristalsis is reflexly set up which reaches the distal part of the large intestine. The obstruction which is presented by the pelvi-rectal flexure is overcome, the sphincter presumably relaxes, and the fæces enter the rectum. Distension of the rectum by the sudden entry of fæces gives rise to a *perineal* sensation, and the desire to defæcate. If this desire is acceded to, a co-ordinated reflex results which empties not only the rectum, but every part of the bowel between the middle of the transverse colon and the anus. The diaphragm descends, the glottis is closed, and the abdominal muscles and levator ani contract. Waves of peristalsis pass over the distal part of the colon; the sphincter ani is relaxed and the fæces are evacuated through the narrow anal canal.

The *sympathetic* inhibits the whole of the large intestine and closes the sphincter ani. The *parasympathetic* innervation is motor and relaxes the sphincter ani. The vagus supplies the proximal colon or cæcum, the sacral autonomic supplies the distal colon and rectum.[1]

FUNCTIONS OF THE LARGE INTESTINE.—The large intestine *secretes* mucus which facilitates the easy passage of the fæces. It passes on the unabsorbed fraction of ingested iron, calcium, and phosphate (pp. 208, 999). *Absorption* of water, salt, and glucose occur in the large intestine.

The Fæces.—The fæces are derived partly from the ingested food, but mainly from the intestinal secretions. The fæces in starving animals are decreased in bulk but differ comparatively little in composition from those of normally fed animals. If vegetables and coarsely ground cereals are excluded from the diet, the fæces have a fairly constant composition, *i.e.* water, 65%; solid material, 35% (weight of *dried* fæces is about 30–50 g.) made up approximately as follows: ash, 15%; ether-soluble substances, 15%; nitrogen, 5%.

The *ash* consists mainly of compound of calcium, phosphate, iron, and

[1] The motility of the large intestine, its vascularity and the rate of mucus secretion are readily modified by the emotional state (see Grace, Wolf, and Wolff, *The Human Colon*, N.Y., 1951.

magnesium. The *ethereal* extract consists of fatty acids, neutral fat, a little lecithin, traces of cholic acid, its decomposition product dyslysin, and coprosterol, (derived from cholesterol). The fæces contain many *desquamated epithelial cells* and *bacteria*, most of which are dead. On a fat intake of 100 g. daily not more than 5–7 g. are normally lost in the fæces; on a protein intake of 100 g. the fæcal N content should not exceed 1·5 g. (corresponding to 10 g. of protein).

Ingested cellulose passes out unchanged, and substances which are enclosed in a cellulose wall escape digestion and absorption. The increased bulk of this undigested residue stimulates intestinal peristalsis. The passage of the food through the bowel is therefore quickened, and the digestive ferments have insufficient time to exert their full action (see Table, *infra*):

Diet.	Moist Fæces. g.	Dried Fæces. g.	Percentage of Ingested Food.	Nitrogen Lost. g.
Bread from fine flour .	133	25	4·0	2·1
Bread from coarse flour	252	41	6·6	3·2
Brown bread . .	317	76	12·0	3·8

As the cellulose content of the food is increased by coarser milling of the flour, the bulk of the fæces is increased. They contain more water and solids; more of the ingested food is undigested and more food nitrogen is lost to the body.

The alterations in the composition of fæces in pancreatic disease and biliary obstruction are described on pp. 794, 798.

DISORDERS OF MOVEMENTS OF ALIMENTARY CANAL

Cardiospasm.—In this condition ingested food is held up at the lower end of the œsophagus owing to the presence there of a "block" of undetermined character. Above the "block" the œsophagus becomes distended (Fig. 538) and its muscle coat hypertrophies. When the contents of the œsophagus form a column 8 inches high, the "block" is overcome, and any excess food or drink above this height enters the stomach.

The "block" has been attributed to *spasm* of the cardiac sphincter (*cardiospasm*), but in fact sphincter tone is not excessive. Alternatively it has been suggested that the cardiac sphincter forms an obstruction because it *fails to relax normally* (*achalasia*) during swallowing. The anatomical lesion in cardiospasm is said to be local degeneration of Auerbach's plexus. Such a lesion would produce cardiospasm only if the vagi were inhibitory to the sphincter, which is probably not the case. The passage of peristaltic waves over the œsophagus depends on the integrity of the vagi. Vagal denervation of the lower end of the œsophagus would prevent the passage of the peristaltic wave and so set up a non-conducting area which would constitute a functional "block." Œsophageal paralysis of this type is produced in dogs by appropriate vagal section (p. 806) (cf. Hirschsprung's disease p. (816).

Gastric Motility in Disease.—Stimulation of any afferent visceral nerve increases the contraction of the pylorus and prevents its proper relaxation. A *gastric ulcer* situated near the pylorus, disease of the *gall-bladder*, and chronic *appendicitis* may all act in the same way, and reflexly via the sympathetic nerves prevent pyloric relaxation and delay the emptying of the stomach. The musculature of the stomach in cases of duodenal ulcer is generally *overactive* so that initially the organ tends to empty itself rapidly.

Hypertrophic pyloric stenosis is a condition which occurs in young infants, usually boys, in whom for some unknown reason the pyloric sphincter is in a state of persistent spasm. This overactivity leads to hypertrophy of the sphincter. Compensatory hypertrophy of the stomach wall develops in order to overcome the obstacle. The passage of the peristaltic waves over the stomach may be visible through the abdominal wall.

Fig. 538.—Œsophagus in Cardiospasm. (Barclay, *Digestive Tract*, Cambridge.)

Note the immense dilatation of the œsophagus as shown radiographically after an opaque meal.

Ileal Stasis.—Reflex spasm of the ileo-colic sphincter may be set up by an inflamed appendix and thus cause delayed emptying of the ileum.

Hirschsprung's Disease.[1]—In this condition, which affects young children, there is degeneration of the cells of Auerbach's plexus in the region of the pelvi-rectal junction; sometimes a greater length of colon may be involved. It is supposed that the denervated region fails to transmit the peristaltic wave and so constitutes a functional "block" (cf. p. 815). The colon above the "block" (*i.e.* the non-conducting region) becomes greatly distended. There is marked constipation, great abdominal distension, and attacks of complete intestinal obstruction may occur. Excision of the affected segment of the colon has yielded satisfactory results.

Constipation.[2]—Two main varieties are recognized:

(1) *Colic Constipation.*—The passage of the fæces through some part of the colon is delayed, but the act of defæcation is normal.

(2) *Dyschezia.*—In this condition the fæces arrive normally at the pelvic colon, but their final evacuation is not adequately performed. If the call to defæcation is neglected, the rectal wall relaxes and the desire passes off. The rectum becomes in time excessively distended (Fig. 539), the contractility of its walls is impaired, and evacuation becomes difficult even when a great effort is made. *Fragmentary* defæcation may take place, *i.e.* small amounts

[1] Bodian *et al.*, *Lancet*, 1949, i, 6.

[2] Hurst. *Constipation and Allied Intestinal Disorders*, London, 2nd edn., 1919.

are passed, but much fæces are left behind in the rectum to continue distending it.

Results of Constipation.[1]—It is still uncertain how the symptoms of constipation are produced. A once popular theory was that of *intestinal auto-intoxication, i.e.* that poisons are elaborated in the digestive tract as a result of constipation, and are absorbed through the mucous membrane and injure the individual. Critical examination of the evidence lends little support to this view. Probably most of the symptoms of constipation are produced *mechanically*. Experiments were performed on healthy men who normally emptied their bowels once or twice daily. They ate an ordinary liberal diet and withheld from defæcation for 90 hours. Each subject developed the symptoms which are regarded as characteristic of "auto-intoxication," *i.e.* foul breath, furred tongue, impaired appetite, flatulence, nausea, loss of power of attention, depression, restlessness, headache, insomnia, and irritability. A barium meal showed that the ileum was empty and that all the food remnants had accumulated in the colon. There was little increase in the urinary indican excretion. An enema was then administered and the bowel evacuated; the distress was promptly relieved, and in an hour or two the subjects felt quite normal again. The symptoms complained of could obviously not have been due to toxic absorption; had "toxæmia" been the cause, a longer time would have had to elapse after evacuation of the bowel before the concentration of toxin in the body could fall sufficiently to produce relief. As Alvarez points out, "a drunken man is not immediately sobered when his whisky bottle is taken from him." The symptoms of constipation are largely due to the *distension and mechanical irritation of the rectum*. Masses of cotton wool packed into the rectum produce identical effects.

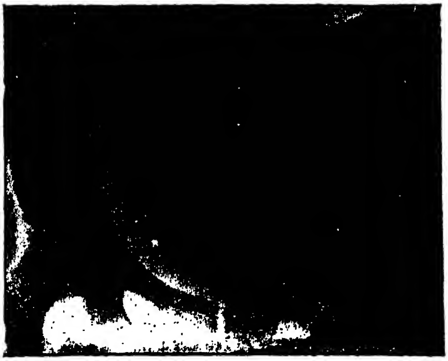


FIG. 539.—Distended Rectal Ampulla seen Radiographically in a case of Dyschezia. (Barclay, *Digestive Tract*, Cambridge.)

It must be remembered that the rectum is very sensitive. A rise of internal pressure of 2–3 mm. Hg is perceptible, and a rise of 20–60 mm. Hg causes much distress. Afferent impulses from the bowel, as is well known, can easily influence the emotions and the mental processes—*e.g.* the occurrence of sleepiness after a meal, or of irritability and lack of concentration when hungry. A patient with a jejunal fistula always went to sleep if the intestine was made to contract actively by the introduction of a small balloon.

The type of patient who is the subject of so-called "auto-intoxication" is usually keenly aware of what goes on in his body and responds strongly to visceral stimuli. Patients of a stolid type may go for weeks or months without

[1] Alvarez, *Physiol. Rev.*, 1924, *4*, 352.

passing a motion and be able to eat well and do a good day's work. A case is recorded of a man who went from June 18, 1900, to June 21, 1901, without a motion. Towards the end of that time he belched a good deal and suffered from some pain; the abdomen was distended; he felt weak and lost some weight. After the colon had been cleaned out (a none-too-simple procedure in this patient) he recovered rapidly.[1]

THE LIVER [2]

Structure of the Liver.—The liver consists of lobes which are subdivided into lobules. The lobule is made up of ramifying columns of *hepatic* cells (Fig. 542); the cell outlines are often indistinct so that the columns form a syncytium. The portal vein, hepatic artery, and bile ducts, surrounded by a connective tissue capsule enter the liver and branch repeatedly in the substance of the organ[3] (Fig. 540). The portal vein divides into branches, the

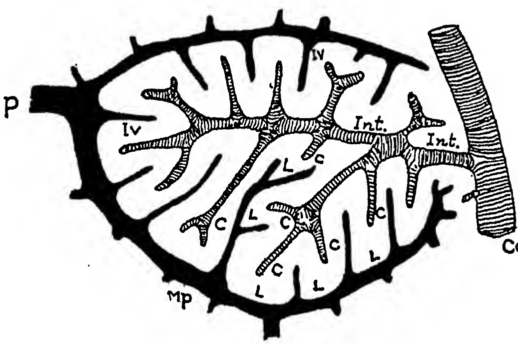


FIG. 540.—Diagram showing Vascular Arrangements in Liver Lobule.

Portal vein branches are shown black: *P*, large branch; *Mp*, medium-size branches; *Iv.*, *inter*lobular veins surrounding lobules.
Hepatic vein branches are shown hatched: *C*, central (*intra*-lobular) vein draining the lobule; *Int.*, *Col.*, successively larger branches carrying blood away from the liver.
White area, *L*, *L*, liver lobules penetrated by vascular capillaries.

interlobular veins, which surround the lobules; from these vessels blood passes between the liver cells in *vascular capillaries* to reach the centre of the lobule where it drains into the *intralobular* branches of the hepatic vein. The hepatic artery likewise divides into branches which accompany those of the portal vein between the lobules; ultimately the hepatic artery blood also

[1] The following anecdote serves as a useful contrast to the foregoing case and shows the wide range of variation in normal people with respect to their alimentary habits. A friend complimented a scholar on his healthy appearance, saying: "Your face is like that of pig-rearers and usurers." The scholar replied: "On my faith, both occupations are forbidden me; but there are twenty-four privies from my lodging-place to the house of study, and whenever I go there I test myself in all of them."

[2] Himsworth, *Lectures on the Liver and its Diseases*, 2nd edn., Oxford, 1950. Lichtman, *Diseases of the Liver, Gallbladder and Bile Ducts*, 2nd edn., London, 1949.

[3] Daniel and Prichard, *J. Physiol.*, 1951, *114*, 521.

enters the vascular capillaries, where it mixes with the blood from the portal vein. The vascular capillaries have no specific endothelial lining but ramify between the hepatic cells which constitute their walls (Figs. 541, 542). The intimate contact between the blood and the liver cells is well demonstrated by injection experiments; the injection material often penetrates from the vascular capillaries into the interior of the liver cells. This is an ideal arrangement, as the liver has to transform, or modify, many of the constituents of the blood. At intervals along the vascular capillaries are the *stellate cells of Kupffer* (Fig. 542) which are part of the *macrophage* or reticulo-endothelial

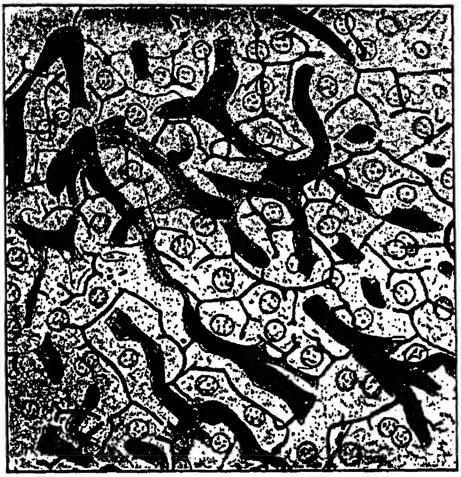


FIG. 541.—Arrangement of Vascular and Bile Capillaries in the Liver. (Schaffer, *Lehrbuch der Histologie.*)

B.C. = bile capillaries (*thin* dark channels).
V.C. = vascular capillaries (*wider*, fainter channels).
N. = nuclei of hepatic cells.
Drawn from a specimen in which both the blood vessels and the bile passages had been injected. The *two networks are quite distinct and separate.*

system (p. 187). They vary in number in different species; in man there are few.

Bile is formed in tiny vacuoles in the interior of the hepatic cells and is discharged through fine canaliculi into the *bile capillaries* (Fig. 541, B.C.). These (like the vascular capillaries) have no specific endothelium, but ramify between, and are lined by, the liver cells. Fig. 541 shows how there is always hepatic cell tissue between the fine bile capillaries and the much wider vascular capillaries, so that normally the blood and the bile are kept apart and *never mix.* The hepatic cells are well placed to transfer materials from the blood into the bile. At the periphery of the lobule the hepatic cells becomes continuous with, and transformed into, the cubical cells lining the *bile ducts.*

STREAMLINE FLOW IN THE PORTAL VEIN.—The portal vein is formed by the union of the splenic and superior mesenteric veins; it can be shown by injecting Indian ink into these veins, that the two entering currents of

blood remain fairly distinct in the portal vein ; the *splenic* blood thus passes mainly to the *left* lobes and the *intestinal* blood mainly to the *right* lobes of the liver. The right lobes consequently receive *directly* a larger share of the absorbed foodstuffs. The blood flow through the liver in man has been estimated to be about 1·5 L per minute.[1]

O_2 Supply of the Liver.—The O_2 saturation of the blood in the hepatic artery is at the usual arterial level of 95% ; in the portal vein (cat) it is about 50% (*i.e.* lower than the resting *mixed* venous blood level) ; in the blood leaving the liver in the hepatic vein it is only 20%. The O_2

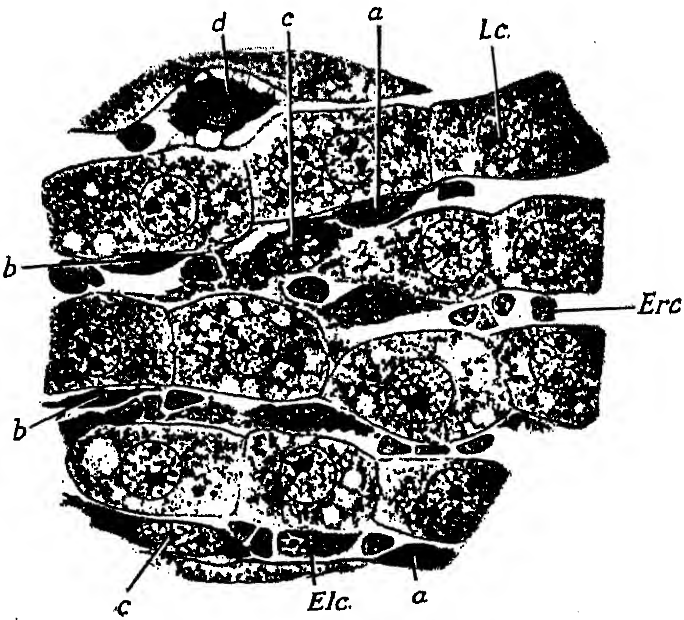


FIG. 542. Section of Liver showing Hepatic Cells, Sinusoids, and Macrophages. (Maximow and Bloom, *Text Book of Histology*, W. B. Saunders.)

Liver of rabbit injected intravenously with Indian ink, which is taken up by the macrophages. *Lc*, liver cell; *Erc.*, erythrocyte in lumen of sinusoid; *Elc.*, eosinophil leucocyte; *a, b, c, d*, macrophages (stellate cells of Kupffer) in various stages of activity *c, d*, macrophages laden with Indian ink.

tension in the vascular capillaries at the *centre* of the lobule is thus *much lower* than in resting tissues elsewhere. The portal vein brings far more blood to the liver than the hepatic artery ; in the cat the ratio is 4 : 1. The ratio of O_2 supplied to the liver cells by the portal vein and hepatic artery (in the cat) is 5 : 3 (8 c.c. of O_2 in all per 100 c.c. of hepatic blood flow).

As the O_2 supply of the cells in the centre of the lobule (*centralobular cells*) is only *just adequate*, any condition which reduces it (or calls for an increase in the O_2 supply) is liable to cause great damage to these cells.

[1] Bradley *et al.*, *J. clin. Investig.*, 1945, *24*, 890.

CAUSES OF HEPATIC ANOXIA.—Centralobular damage or necrosis resulting from an inadequate O_2 supply to the liver may occur in the following conditions:

(i) *Diminished blood flow* through the liver, *e.g.* (*a*) when the cardiac output is decreased (in "shock," hæmorrhage, heart failure); or (*b*) in partial obstruction of the portal vein or of the hepatic artery.

(ii) Breathing air containing O_2 *at a lowered tension* (*e.g.* at high altitudes).

(iii) When the *metabolism* of the liver is *raised*, *e.g.* in hyperthyroidism; the O_2 needs of the liver are increased and the previously adequate O_2 supply becomes insufficient. Hepatic insufficiency is frequently found in hyperthyroidism (p. 991).

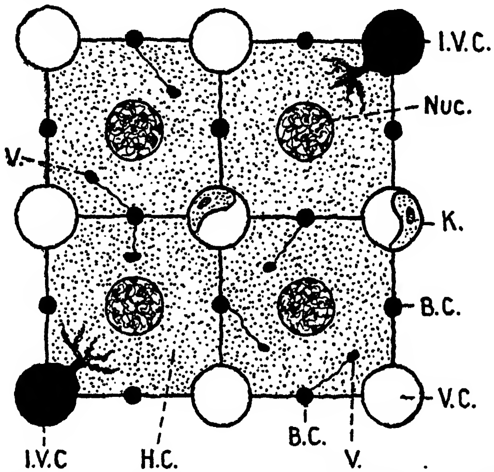


FIG. 543.—Diagram of Liver Structure. (Modified from Merkel and Jordan.)

Four adjacent liver cells are shown.
V.C. = vascular capillary. B.C. = bile capillary. Nuc. = nucleus of liver cell.
K. = Kupffer cell. V. = bile vacuole in liver cell.
I.V.C. = appearance of injected vascular capillary, showing how injection material penetrates into the substance of the liver cell.
Note how both the vascular and the bile capillaries are surrounded by liver cells.

(iv) When the liver cells become *swollen*; as a result the vascular capillaries are compressed and obstructed so that the blood supply of the centralobular cells is reduced. Swelling (œdema) of the hepatic cells is a regular initial effect of any injury to the liver; this explains why poisons absorbed from the intestine produce lesions which are usually most marked not at the periphery but in the *centre* of the lobule, in spite of the fact that the peripheral cells receive the highest concentration of the poison via the portal radicles. Swelling of the hepatic cells associated with centralobular necrosis also occurs when the cells become extensively laden with *fat* (so-called *fatty liver* or *fatty infiltration*, p. 867).

BLOOD STORAGE IN THE LIVER.—The calibre of the hepatic vascular capillaries may vary considerably during life; the capillaries thus serve as a storage area for blood which can be discharged when necessary into the general circulation (cf. spleen, p. 226).

Portal Obstruction.—This may occur in a variety of circumstances:

(i) (*a*) *Ligature of one branch* of the portal vein produces atrophy in the corresponding ischæmic area of the liver; the hepatic cells almost completely disappear, but the bile ducts, blood vessels, Kupffer cells, and connective tissue are more resistant, and survive. The unaffected lobes (which are still normally vascularized) undergo compensatory hypertrophy, up to five times their original weight.

(*b*) Ligature of the *main* portal vein in animals produces marked atrophy of the liver; if part of the liver is then excised *no regeneration of the remaining tissue takes place* (p. 826).

(*c*) Occlusion or narrowing of the main portal vein or of one of its branches may also occur clinically.

(ii) Obstruction to the portal flow is a common initial result of serious hepatic injury (*hepatitis*) owing to œdema of the cells. Later when hepatic cells die they are replaced by fibrous tissue which contracts down and obliterates the contained blood vessels; the more extensive the fibrosed area, the more severe is the resulting interference with the portal venous flow.

CLINICAL RESULTS OF PORTAL OBSTRUCTION FROM ANY CAUSE.—The main results are as follows:

(i) There is a *rise of portal vein pressure*; direct measurements at operation in normal man give values of 15–22 cm. H_2O (compared with 5–12 cm. H_2O in the ankle veins); in hepatic fibrosis (so-called cirrhosis of the liver), the portal venous pressure may be as high as 40 cm. H_2O.

(ii) *Ascites* (an abnormal accumulation of fluid in the peritoneal cavity) develops. The factors which contribute to the production of ascites are as follows:

(*a*) Owing to the raised pressure in the portal vein there is a corresponding rise of *capillary blood pressure* in the abdominal viscera leading to an excessive outflow of fluid into the peritoneal cavity (p. 18).

(*b*) In many patients with portal obstruction the hepatic cells are also damaged leading to a *fall of serum albumin* concentration; this further facilitates the escape of fluid from the capillaries. The protein content of the ascitic fluid varies widely, *e.g.* from 0·1 to 3·8%; this finding indicates that the *permeability* of the capillary wall to protein may be increased to a variable degree.

Ascites may develop in cases with disease of the liver parenchyma even without much rise of portal pressure if the plasma protein concentration is markedly reduced, *e.g.* to about half normal. Fig. 544 shows that the greater the fall in serum albumin the more likely is ascites to develop. Blood transfusion may cause removal of the ascitic fluid presumably because it raises the plasma protein level.

(*c*) For some unknown reason ascites may be induced by a diet rich in *meat*; the fluid is absorbed when meat is excluded from the diet. The noxious agent in meat is not its protein content, because protein-free meat-extracts are even more effective in producing ascites than is an equivalent amount of whole meat.

(*d*) In some cases of ascites an *antidiuretic agent* is present in the urine; this finding is interpreted as indicating that the pituitary antidiuretic hormone is present in the blood in excessive amounts, probably because the damaged liver cannot inactivate it in the normal way. As a result there is excessive

reabsorption of water by the kidney leading to an increase in the volume of the body fluids; much of this fluid accumulates in the peritoneal cavity because of the abnormalities described in (*a*) and (*b*) above.

(*e*) In some patients with chronic hepatitis (cirrhosis of the liver) the ascitic fluid accumulates very rapidly and may have to be withdrawn from the peritoneal cavity every few weeks. An important contributory factor is the failure of the kidney to excrete Na^+ ions; Na^+ is retained, leading secondarily to water retention.[1] On a liberal salt intake the amount of Na^+ excreted in the urine may be less than 1 m.Eq. (0·023 g.) daily. If the Na^+ intake is drastically reduced the ascites tends to clear up. It is suggested that there is *hypersecretion of adrenal corticoids* which cause an excessive degree of

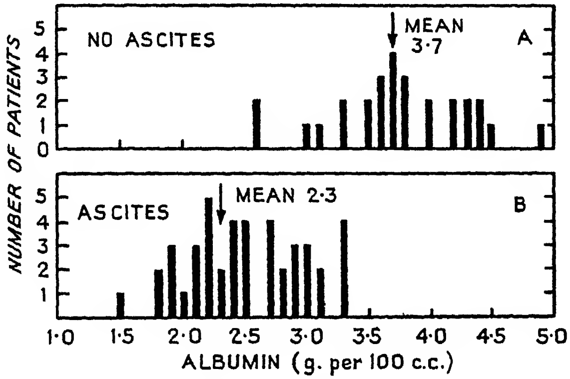


FIG. 544. Relation of Serum Albumin Concentration to Ascites in Cirrhosis of Liver (Chronic Hepatitis). (Post and Palik, *Arch. int. Med.*, 1942, *69*).

Note that the mean value for serum albumin in cases with ascites was 2·3%, and in cases without ascites 3·7%.

reabsorption of Na^+ from the lumen of the renal tubules into the blood. It is noteworthy that the Na^+ content of the sweat and saliva is also low in these patients (p. 945).

Thus in patients with portal obstruction ascites may be due to: increased capillary pressure in the portal bed; reduced serum albumin concentration; a toxic factor in meat; excess antidiuretic hormone: sodium retention resulting perhaps from excess adrenal corticoids. It is clear that hepatic and renal dysfunction are important contributory factors.[2]

(iii) A *collateral circulation* is established between the portal and the systemic circulation, *e.g.* at the junction of œsophagus and stomach, in the mucosa of the anus, and in the abdominal wall. *Hæmorrhages* may occur from the dilated venules of the bowel.

Functions of the Liver.—The functions are very numerous and are discussed in various sections of this book. They are summarized below, where adequate cross-references are given.

[1] Eisenmenger *et al.*, *J. clin. Investig.*, 1950, *29*, 1491.

[2] Ascites may also occur: (1) in *cardiac failure*: it is due to raised intra-abdominal venous pressure and renal dysfunction (cf. cardiac œdema, p. 111); (2) in *nephrosis*: it is due to lowered serum albumin concentration (cf. p. 113).

(1) STORAGE ORGAN.—The liver stores glycogen (p. 837), fat (p. 866), probably proteins (p. 138), vitamins, *e.g.* vitamin-*A*, antipernicious anæmia factor (p. 197), other substances concerned in blood formation and regeneration (p. 210), and blood (p. 821).

(2) SYNTHESIS.—The liver synthesizes the plasma proteins (p. 137), fibrinogen (p. 140), prothrombin (p. 142), and (by virtue of its mast cells) heparin (p. 145).

(3) BILE SECRETION (p. 797).

(4) FORMATION (p. 160) and DESTRUCTION OF RED CELLS (p. 187); relation to JAUNDICE (p. 190).

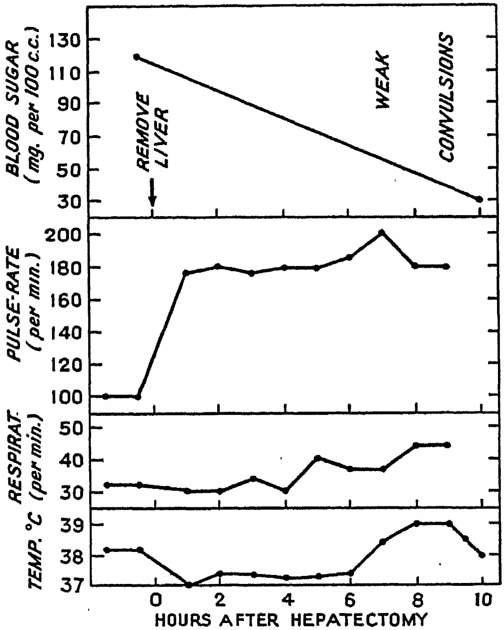


FIG. 545.—Changes in Blood Sugar, Pulse Rate, Respiration, and Temperature, following Removal of Liver in Dog. (Modified from data by Mann and Magath, *Arch. int. Med.*, 1922, *30*, 73.)

The animal died 10 hours after hepatectomy.

(5) DETOXICATING FUNCTION (p. 828).

(6) METABOLISM.—The liver is pre-eminently the central organ of metabolism. To discuss its rôle adequately would mean considering in detail the metabolism of carbohydrate, fat, and protein. It is more convenient, therefore, to consider the rôle of the liver when dealing with the metabolism of the individual foodstuffs. The following sections should be consulted:

(i) Carbohydrate Metabolism (pp. 837 *et seq.*), especially Rôle of Liver Glycogen (p. 838) and Regulation of Blood Sugar (p. 855).

(ii) Fat Metabolism (pp. 863 *et seq.*), especially discussion on pp. 866 *et seq.*

(iii) Protein Metabolism (pp. 876 to 904).

Complete Extirpation of Liver.—The liver can be completely extirpated in animals. Much has been learnt about the functions of the liver from such studies. The animal recovers rapidly from the operation and to all outward appearances seems fairly normal for a few hours. The changes which develop in pulse rate, respiration, and temperature are shown in Fig. 545. The principal results of complete hepatectomy are as follows:

(1) There is a progressive fall of blood sugar (development of *hypoglycæmia*), *e.g.* to 40 mg–% or less, producing the characteristic symptoms for the species; in the dog there is marked muscular weakness, followed by convulsions, coma, and death. The aggravation of the clinical state runs parallel with the fall of the blood sugar. The somewhat different symptoms of hypoglycæmia in man are discussed on p. 915.

The administration of glucose in doses of 0·25–0·5 g. per kg. body weight to a flaccid and comatose animal produces astonishing effects: within 30 seconds the dog can walk and respond to a call; the heart beats more strongly. The blood sugar, immediately following the injection, reaches a high level; when it decreases again symptoms reappear. The development of fatal hypoglycæmia after hepatectomy proves that the liver is the organ principally *responsible for the maintenance of the normal level of the blood sugar* (cf. p. 856).

Light is thrown on the fate of certain sugars in the body by the following observations in the hypoglycæmic, liverless animal. *Maltose, mannose, fructose* and *glycogen* act, on intravenous injection, like glucose; these four substances can thus be converted somewhere outside the liver into glucose. *Galactose,* however, is neither converted into glucose nor used in the liverless animal. This last observation is the basis of the *galactose test of liver efficiency* (p. 830).

(2) There is a progressive fall of the *blood urea* and rise of the *blood amino-acids* (p. 886, Figs. 566, 867). This proves that the liver is the chief site of deamination of amino-acids and the *only organ in which urea is formed.*

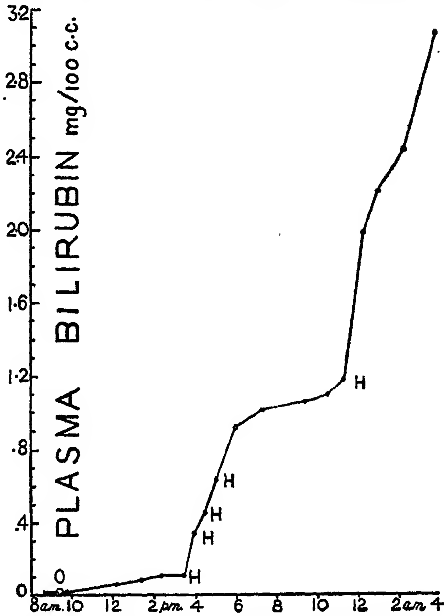


FIG. 546.—Accumulation of Bilirubin in the Blood Following Removal of the Liver; Conversion of Hæmoglobin into Bilirubin. (Redrawn from Mann *et al.*, *Amer. J. Physiol.*, 1924, *69*, 399).

Experiment on dog.
Ordinate: plasma bilirubin in mg–%.
At O, total removal of liver. Note gradual rise in bilirubin level. At each point marked H injection of hæmoglobin intravenously. Note resulting marked increase in bilirubin level indicating extra-hepatic bilirubin formation. Considerable amounts of bilirubin were also excreted in the urine.

(3) *Jaundice* develops after total hepatectomy; bilirubin accumulates in the blood and tissues (Fig. 546), and is excreted in the urine. Jaundice develops constantly in animals which survive longer than 6 hours; the plasma and fatty tissues become coloured. The blood gives a positive Van den Bergh reaction for bilirubin (p. 189), first an indirect and than a biphasic reaction. The removal of the spleen and of all the abdominal viscera does not alter the rate of development of the jaundice. If hæmoglobin is injected intravenously into the liverless animal considerable additional amounts of bilirubin are formed (Fig. 546). These observations prove that *bilirubin can be formed extra-abdominally* (cf. p. 188). The liver, however, is the only organ normally concerned with the *execretion* of bile pigment.

(4) The *coagulability* of the blood is depressed owing mainly to a marked decrease in the concentration of plasma *prothrombin* and perhaps also to a decrease in plasma *fibrinogen.*

(5) *Hepatic insufficiency* develops. If the blood sugar is kept normal by the continuous infusion of glucose the animal may survive for 18–24 hours. Finally, however, restlessness, dyspnœa, and vomiting occur; the animal becomes ataxic, and does not appear to hear or see. Coma and anuria develop, and death takes place quite suddenly. These symptoms are unrelieved by glucose and are not due to hypoglycæmia; these terminal manifestations are due to loss of some unknown liver functions (*i.e.* to some undefined form of hepatic failure). (Cf. *acute clinical hepatic failure* (p. 833).)

Partial Extirpation of the Liver.[1]—The results produced depend on (i) the physiological reserve of liver tissue; (ii) the regenerative power of the liver; (iii) its nutritional state.

(i) The *physiological reserve* is indicated by the fact that the body contains liver tissue far in excess of the minimal amount necessary for normal physiological function. Thus, in the dog, after removal of 80% of the liver, bile salts and bile pigments are not retained in the blood or excreted in the urine. Even if 90% of the bile ducts are ligated the volume of bile secreted remains normal.

(ii) The *regenerative power* of the liver is illustrated by the following observations. In the dog, if three-fourths of the liver is removed, proliferation takes place in the remaining tissue as a result of active mitotic division of the cells; the original liver mass is restored in 6–8 weeks.[2] The bile ducts at the periphery of the lobules also sprout and bud off new clumps of cells; blood vessels and connective tissue soon invade the newly formed areas. Excision can be repeated many times and is always followed by regeneration. Regenerative processes play an important part in the repair of the liver following the administration of hepatic poisons.

(iii) The importance of the *nutritional state* and the *blood supply* are considered on pp. 822, 827.

Chronic hepatic insufficiency may be produced as follows: Establish an Eck fistula, *i.e.* a lateral anastomosis is established between the portal vein and the inferior vena cava; the portal vein is tied headwards to the stoma, thus cutting off most of the blood supply to the liver. It becomes reduced to half or less of its original size; within 2 months remove 60% of the liver. No regeneration now takes place, so that less than one-fifth of the original amount of liver tissue is left and very few of the cells which are present appear normal on histological examination. In spite of this, the animal maintains *fairly normal health.* The chief abnormalities present are as follows: (i) the blood sugar tends to be slightly below normal; (ii) after injection of insulin the recovery of the blood sugar is greatly delayed (Fig. 582); (iii) pancreatectomy produces only a slight hyperglycæmia as hepatic glucogenesis is impaired; (iv) adrenaline produces a less marked rise in blood sugar; (v) poisons are not well tolerated.[3]

Chemical Composition of the Liver.—The Table on p. 827 shows how the weight and chemical composition of the liver may be influenced by alterations in the composition of the diet (data of Bollman and Mann for the *dog*).

The lowest and highest percentage in each case are shown in *italics.*

[1] Bollmann and Mann, *Ergeb. Physiol.*, 1936, *38*, 445.

[2] Regeneration must have taken place very actively in the case of Prometheus, whose liver was gnawed daily by a tormenting eagle.

[3] The newer clinical tests of hepatic efficiency were not carried out on these animals.

Liver weight may vary 4-fold; fat content per cent., 32-fold; glycogen content per cent., 150-fold; and water content, 2-fold.

Diet.	Liver Weight as per cent. of body weight.		Fat as per cent. of liver weight.		Glycogen as per cent. of liver weight.		Water as per cent. of liver weight.	
	Low.	High.	Low.	High.	Low.	High.	Low.	High.
Mixed . .	*1·8*	3·3	*1·6*	5·2	0·8	7·1	68·4	*74·6*
High protein .	3·1	4·4	2·3	3·1	2·3	5·8	70·2	72·7
High carbohydrate	4·4	4·6	1·7	2·3	7·1	*11·0*	71·8	73·2
High fat . .	4·1	*7·5*	23·4	*52·2*	*0·07*	2·3	*35·6*	57·4
Fasting . .	2·5	3·4	5·2	15·7	0·3	2·1	65·8	73·6

Relation of Composition of Diet to Liver Function.—There are a number of important observations showing that the protein, carbohydrate and fat intake are related to the efficient functioning of the liver.

(1) Effects of High Fat Diet and Rôle of Lipotropins (p. 867).

(2) Protective Action of Methionine.—As pointed out on p. 867, methionine, by donating CH_3 groups to form choline, is an indirect lipotropin. In addition it has been found that methionine (and to a smaller extent cysteine) *protects the liver* against the noxious action of certain poisons and especially against chloroform. A dog which has been kept on a low protein diet for a long time is highly susceptible to chloroform poisoning; *light* anæsthesia for 20 minutes kills the animal; post-mortem, extensive necrosis of the liver cells is found. But if the dog is given methionine (or cysteine) a few hours previously, it can withstand *deep* anæsthesia for 40 minutes with little subsequent clinical disturbance and no signs of liver injury. A large meat meal (or plasma protein injected intravenously) has a similar protective action, doubtless because of its S-containing amino-acid content. (It should be noted that choline is not protective against chloroform poisoning.)

(3) Effects of Grave Protein Insufficiency.—(i) Under certain conditions, as yet imperfectly determined, grave protein insufficiency may, after a delay of some weeks, produce, in rats, acute hepatic insufficiency due to *massive necrosis* of the liver. The necrosis involves large areas which are separated from each other by normal regions. In the affected areas the liver cells are dead and there is much congestion and many hæmorrhages. If the animal recovers, *post-necrotic* scarring occurs: the normal regions hypertrophy and between them are found scars which interfere with the hepatic blood supply. Special stress has been laid on the fact that the *cysteine* content of these diets is deficient. The experiment makes it clear that in the absence of adequate supplies of amino-acids (needed presumably for the maintenance of the liver proteins and enzymes), the liver cells may die.

(ii) A similar condition of massive necrosis occurs in animals fed on grain grown in soil rich in *selenium*. The selenium in the grain replaces the *sulphur* in the sulphur-containing amino-acids, resulting in hepatic necrosis; the hepatic necrosis is prevented by administering methionine or cysteine.

(4) In general, it may be emphasized that the liver resists many forms of stress best when its stores of carbohydrate and protein are ample; its efficiency is impaired when it is laden with fat. Thus a dose of carbon tetrachloride which is fatal (in the dog) within 24 hours if the liver is fatty at the time of administration is without effect if the liver is well filled with glycogen. In the case of tetrachlorethane, 1 c.c. produces severe intoxication and coma of 6–8 hours' duration if given 24 hours after the last meal; if given 12 hours after food, only mild symptoms result. The toxic action of ethyl alcohol on the liver varies inversely as its glycogen content; the toxic action of trinitrotoluene and arsphenamine is potentiated by protein deficiency and a high fat intake.

TOXIPATHIC AND TROPHOPATHIC HEPATITIS.—Himsworth recognizes two forms of hepatitis (liver damage):

(i) *Toxipathic*, due to the action of poisons (such as those mentioned in (4) above).

(ii) *Trophopathic*, due to nutritional disturbances, *e.g.* protein deficiency and the factors causing "fatty liver" (p. 867). It must also be remembered (*a*) that nutritional disturbances make the liver more vulnerable to attack by poisons; and (*b*) that a diet which is normally adequate may become insufficient in periods of *growth* or during *pregnancy*.

Massive necrosis (also called "acute yellow atrophy") occurs in man (cf. p. 833), but the cause is unknown.

"Detoxicating" and Protective Action of the Liver.—The liver exerts its protective action in a variety of ways:

(1) By CONJUGATION,[1] *i.e.* by combining the unwanted substance (or a derivative of it) with another molecule or chemical group, the resulting compound being excreted in the urine. It does not always follow that the excreted compound is less toxic than the original, as measured by the usual toxicity tests—it may even be *more* toxic. For this reason, and for the additional reason that many normal physiological compounds are chemically "manipulated" by the liver in this way before being excreted, the term "detoxication" commonly applied to conjugation in the liver is a misnomer, and would be better replaced by "protective synthesis." Examples are:

(i) *Conjugation with Sulphate.*—Many phenolic compounds are conjugated in the liver with sulphate, and excreted as "ethereal sulphates" (sulphate esters) (p. 900).

(ii) *Conjugation with Glycine.*—Many aromatic acids that cannot be catabolized in the body are combined with glycine (or other amino-acids in special cases) before excretion. Benzoic acid ($C_6H_5.COOH$) is transformed by the liver into hippuric acid (benzoylglycine, $C_6H_5.CO.NH.CH_2.COOH$); a similar detoxication of benzoic acid also occurs in the kidney (p. 830). On the other hand, phenyl-acetic acid ($C_6H_5.CH_2.COOH$) is "detoxicated" by combination with glutamine (the amide of glutamic acid).

(iii) *Conjugation with Glucuronic Acid.*—Many drugs and hormones containing OH groups (either alcoholic or phenolic) combine with glucuronic acid to form *glucuronides*; thus pregnanediol (from progesterone) is excreted in the urine as pregnanediol glucuronide after conjugation in the liver, and similarly for the other urinary steroids (pp. 1078, 1080).

(iv) *Conjugation with Acetic Acid.*—Aromatic amino compounds react

[1] Williams, *Detoxication Mechanisms*, London, 1947.

in the liver with acetic acid (as reactive "2C fragments" (p. 871)) to form the corresponding acetyl derivatives, *e.g.* sulphanilamide forms acetyl-sulphanilamide, which is then excreted (p. 873).

The mode of treatment of a foreign substance by the liver depends chiefly upon its chemical structure, but partially on the species of animal involved.

(2) By Complete Destruction.—Many compounds foreign to the body are destroyed in the liver by *complete oxidation.* Example are the alkaloids strychnine and nicotine, and the anæsthetic Na pentobarbital. *Partial oxidation* (or, less often, *reduction*) may precede the conjugation reactions described in (1) above.

Clinical Hepatic Failure.—Disease of the liver in man may result in:

(1) *Excretory Failure.*—Failure to excrete bile leads to retention of the constituents of the bile (p. 804) and an increase in the concentrations of bile

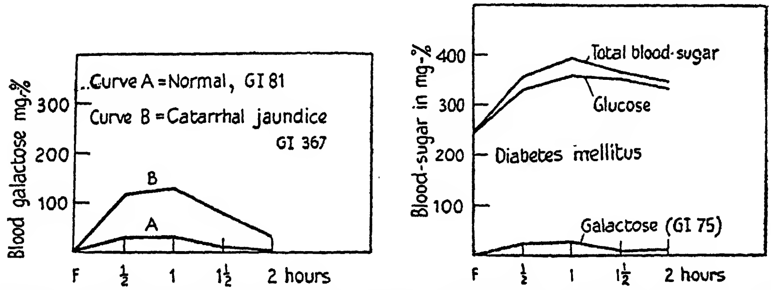


Fig. 547.—Blood Galactose Curves in Hepatic Insufficiency. (McLagan, *Quart. J. Med.* 1940, *9*, 156.)

Left-hand Figure: ordinate: blood galactose in mg-%. Ingest 40 g. of galactose by mouth.
A: normal curve. GI: galactose index=81.
B: abnormal curve (excessive rise of blood galactose) in case of catarrhal jaundice with hepatic insufficiency GI=367.
Right-hand Figure: ordinate: total blood "sugar," blood glucose, or blood galactose in mg-%.
Case of diabetes mellitus. Simultaneous administration of galactose and glucose by mouth. Norma blood *galactose* curve, GI=75; characteristic *diabetic blood glucose* curve.

pigment, bile salts, cholesterol, and alkaline phosphatase in the blood. Bile salts and bile pigments appear in the urine; jaundice develops. The absence of bile from the intestine impairs digestion and, in addition, decreases the absorption of fat and fat-soluble vitamins.

(2) *Portal Obstruction.*—This condition is characterized by development of collaterals between the portal and systemic veins, and ascites (p. 822).

(3) *Parenchymal Failure.*—In this condition excretory failure may or may not be present; the term parenchymal failure is used to indicate that the liver cells are failing to carry out their other functions, to a greater or lesser degree.

A number of laboratory tests are available for assessing the form and extent of *parenchymal* hepatic failure.

Laboratory Tests of Liver Function.[1]—The facts set out on p. 826 about the difficulty of producing experimental liver insufficiency, without completely removing the liver, should be borne in mind when considering means of testing liver function clinically. Extensive liver damage must be present in man before obvious signs of insufficiency present themselves. Some liver

[1] Gray, *Quart. J. Med.*, 1950, *19*, 263.

functions are, however, impaired sooner than others and these may serve as more sensitive tests. In disease of the liver *parenchyma* the excretion of bile is often not markedly affected and consequently there may be little jaundice and no increase in blood bilirubin; urea formation is another very resistant function. The following tests are of value.

(1) GALACTOSE TOLERANCE TEST.—When galactose is absorbed from the intestine it is normally converted by the liver into glycogen and subsequently into glucose which is dissimilated by the tissues. In the liverless animal, however, galactose is scarcely utilized at all and is only disposed of by excretion in the urine. The test consists of administering 40 g. of galactose by mouth and determining the *blood galactose* at 0·5, 1·0, 1·5, and 2 hours. In normal subjects the blood galactose rises very slightly; in the presence of liver damage a greater rise takes place (Fig. 547). The sum of the four galactose values (at the times stated) in mg–% gives the *galactose index* (Fig. 548). The normal average is 68 and the maximum normal 160. Higher

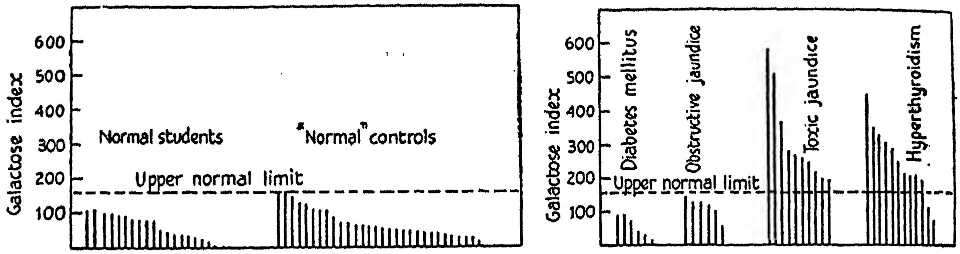


FIG. 548.—Galactose Index in Normal Subjects and in various Clinical States. (McLagan, *Quart. J. Med.*, 1940, *9*, 156.)

Left-hand Figure: normal distribution and upper normal limit of galactose index.
Right-hand Figure: normal galactose index in diabetes mellitus and group of cases with obstructive jaundice; abnormally raised index in toxic jaundice and in hyperthyroidism.

values indicate hepatic insufficiency. The galactose index is not raised in cases of diabetes mellitus (Fig. 547, right G.I.=75), even when the glucose tolerance is markedly impaired. Fig. 548 shows a raised galactose index in toxic jaundice and in hyperthyroidism (in which liver impairment is common, p. 991).

(2) HIPPURIC ACID TEST.—When sodium benzoate is taken by mouth it is conjugated in the liver and kidneys *with glycine* to form hippuric acid. Clinically 6 g. of Na-benzoate are ingested; normally 3–3·5 g. of benzoic acid are excreted in the urine as hippuric acid in the course of the next 4 hours; an excretion of 2·7 g. is taken as the low limit of normality. In liver disease hippuric acid excretion is depressed owing to diminished conjugation.

Renal disease must, of course, be excluded as impaired kidney function gives the same result owing to diminished hippuric acid synthesis and diminished renal excretion of that which is formed.

(3) BROMSULPHTHALEIN EXCRETION TEST.—Normal liver cells secrete this dye (phenol and tetrabromphthalein disodium sulphonate) from the blood into the bile. Five mg. per kg. of the dye are injected intravenously and specimens of blood collected after 5 and 45 minutes. Assume the initial concentration of the dye in the blood to be 100%. After 5 minutes the blood

concentration normally falls to 85%, and after 45 minutes to 5% A value at 45 minutes exceeding 10% indicates liver damage.

(4) Coagulability of the Blood (*Prothrombin Response to Vitamin-K*).—In liver disease the coagulability of the blood is decreased primarily owing to the lowered plasma prothrombin; if there is also a fall of plasma fibrinogen, it constitutes an additional contributory factor; as a result hæmorrhages may occur. In such cases the administration of vitamin-*K* by any route does

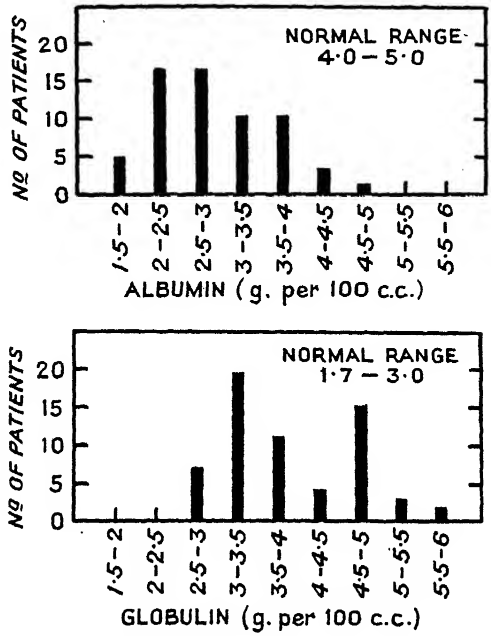


Fig. 549.—Plasma Protein changes in Liver Disease.
(Post and Palik, *Arch. int. Med.*, 1942, *69.*)

Frequency distribution of values (on admission to hospital) for serum albumin and serum globulin in 61 patients with cirrhosis of liver. Note the decrease in serum albumin and the increase in serum globulin.

not raise the abnormally low prothrombin level, *i.e.* the prolonged prothrombin clotting time, characteristic of liver disease is unaffected (p. 153).

(5) Serum Albumin Concentration.—This is lowered in liver insufficiency because the liver is the sole site of serum albumin formation. For some unknown reason serum globulin concentration is raised, often markedly (Figs. 549, 550) (cf. p. 833). High protein feeding fails to restore the serum albumin level as the mechanisms concerned in the manufacture of albumin are impaired.

(6) Empirical (Flocculation) Tests of Liver Function.—The objects of these tests are:

(i) To distinguish between jaundice due to liver *cell damage* (*hepatitis*) combined with inflammatory swelling of the small bile ducts, which gives

a *positive* result, and jaundice, caused by gross obstruction of the larger ducts (by growth or stone) which gives a *negative* result (cf. p. 191).

(ii) To give a *quantitative* measure of the hepatitis present and of changes in the intensity of the process.

In hepatitis there is a characteristic alteration in the distribution of the plasma proteins. Not only is there a fall in serum albumin and a rise in serum globulin (as mentioned above), but the relative magnitudes of the different components of the globulin fraction (α, β, ϕ, γ) are changed. Fig. 550 illustrates

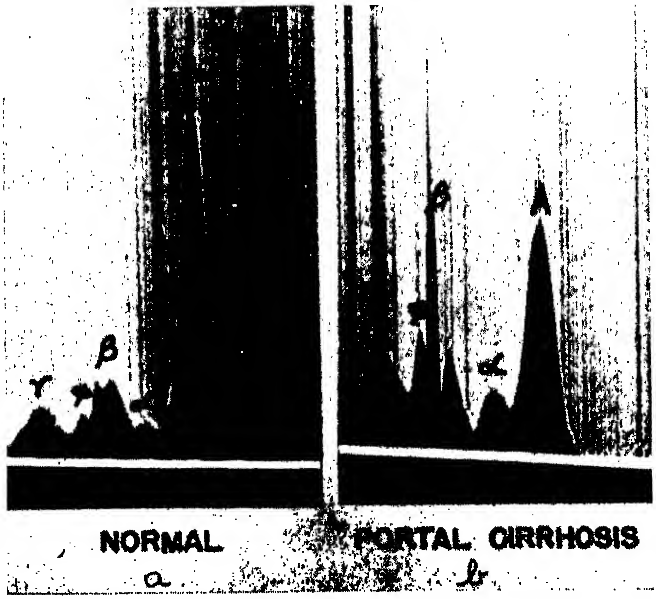


FIG. 550.—Electrophoretic Analysis of Distribution of Plasma Proteins in Patient with Chronic Hepatitis (Cirrhosis of Liver) compared with Normal Subject. (Whitman *et al.*, *J. lab. clin. Med.*, 1950, *35*, 172.) (Cf. Fig. 73).
A = albumin. α, β, ϕ, γ = globulin fractions.

characteristic serum protein changes in a patient with hepatitis. The results of the empirical tests are related to these plasma changes. Various reagents are added to the serum under investigation. When the test is *positive* the added reagent is precipitated out of solution (*flocculation*) giving rise to *turbidity*. The factor mainly responsible for causing the flocculation is the increase in the γ-globulin which is commonly associated with hepatitis though an increase in other globulins may play a minor part; the serum albumin *inhibits* flocculation. The thymol reagent (*infra*), however, is also readily precipitated by *lipids*; a misleading ("false") positive thymol reaction may thus be obtained in cases of nephritis associated with lipæmia. When the test is *negative*, no flocculation occurs.

Three types of reagents are used:

(i) Salts of *divalent metals*, *e.g.* cadmium and zinc sulphates. (These

are used in a concentration much below that required to give complete protein precipitation.)

(ii) A saturated *thymol* solution.

(iii) *Negatively charged colloids:* colloidal gold or a colloidal suspension made by boiling an ethereal solution of cephalin and cholesterol with water. (The γ-globulin does not act on these reagents by neutralizing the negatively charged colloid since it is itself negatively charged.)

All these reagents are unchanged by normal blood sera except the thymol solution which gives a faint cloud; in hepatitis the thymol cloud may be 10–30 times as intense as normal, and flocculation occurs. Hence the thymol turbidity test is specially useful as a quantitative reaction.

The following conditions give *positive* results with most of the tests:

(*a*) *Hepatitis*: positive tests are obtained in over 90% of cases.

(*b*) *Virus* diseases: usually in lymphogranuloma inguinale and sometimes in infective mononucleosis.

(*c*) *Protozoal* disease: usually in Leishmaniasis and sometimes in malaria.

(*d*) Severe *infections*: tests are sometimes positive in such conditions as tuberculosis, pneumonia, infective endocarditis.

(*e*) Tests are sometimes positive in multiple myelomatosis, sarcoidosis and rheumatoid arthritis.

In obstructive jaundice due to pressure on the *larger* ducts by growth or stone, the tests are almost always negative for the first four weeks and maybe for much longer. If, after this time, the tests become positive, it may be inferred that the back pressure of the bile is damaging the liver cells or that infection has supervened (cf. (*a*) and (*d*) above). Thus, a case of jaundice of only a few weeks' duration which gives positive flocculation tests is probably due to hepatitis; jaundice of similar duration giving negative tests is probably due to gross obstruction of larger ducts. If a positive test is obtained in the *absence* of jaundice, and conditions (*b*) to (*e*) (above) can be excluded, hepatitis is probably present (*e.g.* hepatitis due to amœbic infection). If the patient is known to have hepatitis, the condition must be regarded as incompletely healed until the thymol turbidity is quantitatively normal.

The source of the γ-globulins which are responsible for the positive reactions, is obscure. It has been suggested that they are the products of abnormal activity by damaged liver cells; but their occurrence in many other diseases in which the liver is not known to be involved, indicates that they may be derived from other sources, *e.g.* the reticulo-endothelial system, lymphocytes or plasma cells.

Acute (Rapidly Fatal) Hepatic Failure. (*Acute Yellow Atrophy, Acute Massive Hepatic Necrosis*).—The cause of many of these cases is still undetermined; the condition constitutes the nearest equivalent in man to that found in animals after complete extirpation of the liver.

The initial symptoms are indefinite; the patient feels ill and complains of loss of appetite (*anorexia*) and nausea. He is apathetic and spends most of the day in a drowsy state. At night his sleep is disturbed by unpleasant dreams. There is achlorhydria which is resistant to injection of histamine (p. 783). The breath is fœtid, "like the smell of a freshly opened corpse." The patient finally takes to his bed owing to extreme weakness. Vomiting develops which becomes persistent; *jaundice* appears which soon becomes severe. The serum bilirubin rises (*e.g.* to over 20 mg–%; the icterus index may

exceed 100 units) [1]; the blood prothrombin falls (*e.g.* to 10% of normal) as does the plasma fibrinogen (*e.g.* to 0·1 g–%); *bleeding* may occur into the skin and mucous membranes. The blood *urea* is low (*e.g.* 8 mg–%); the blood amino-acid level is raised and amino-acids appear in the urine; the blood NH_3 is raised. The pulse, temperature, and respiration are unaffected until late in the disease when the pulse accelerates and the temperature rises.

The final symptoms are nervous and mental and are perhaps due in part to progressive hypoglycæmia. "Consciousness is clouded," and there are hallucinations and delirium. A rigidity like that found in Parkinsonism develops and there are fine muscular twitchings, but convulsions are rare; finally coma sets in which deepens into death.

The non-specific hepatic tests (*i.e.* the various flocculation reactions) become positive after the first few days and are indicative of "active damage" to the liver cells.

In some cases (*hepato-renal syndrome*) additional signs of renal insufficiency are prominent, *e.g.* decreased volume of urine, retention of nitrogenous substances in the blood, and finally even complete anuria.

[1] *Icterus index.*—This is a measure of the concentration of serum bilirubin and biliverdin, obtained by comparison of the intensity of the yellow colour of the serum with a standard potassium dichromate solution. The index is expressed in standard, though arbitrary, units, and normally lies between 3 and 6 units; above 15 units, clinical symptoms of jaundice may appear.

VIII

METABOLISM

CARBOHYDRATE METABOLISM.[1] DIGESTION, ABSORPTION, AND STORAGE OF CARBOHYDRATES

Carbohydrates of Food.—The principal carbohydrates in the food are :

(i) *Polysaccharides* ($C_mH_{2n}O_n$), *e.g.* starch $(C_6H_{10}O_5)_x$ in vegetable foods. Cellulose and pectins cannot be digested by the enzymes of the human bowel.

(ii) *Disaccharides* ($C_{12}H_{22}O_{11}$), *e.g.* sucrose (saccharose, cane- or beet-sugar), lactose (from milk), and maltose.

(iii) *Monosaccharides* : (*a*) *Hexoses* ($C_6H_{12}O_6$), *e.g.* glucose (dextrose), and fructose (lævulose) in fruits and vegetables. (Galactose is not ingested as such but is split off from lactose).

(*b*) *Pentoses* ($C_5H_{10}O_5$), not in the free form, but in nucleic acid (p. 878) and in certain polysaccharides, *e.g.* the pentosans of fruits and gums.

The metabolic " pathway " of carbohydrates (*i.e.* the sequence of intermediate compounds and reactions by which metabolic transformations are effected) may be joined by certain food constituents which are not carbohydrates, *e.g.* glycerol.

Digestion.—(i) *Ptyalin* of the saliva can only digest starch after the cellulose envelope has been burst, *e.g.* by cooking. It acts in an alkaline, neutral or faintly acid medium (optimally at *p*H 6·5).[2] Ptyalin digestion can thus continue in the stomach for about half an hour until it is arrested by the excessive acidity of the gastric contents. The (unimportant) digestive action of ptyalin may thus be summarized :

$$\text{Starch} \longrightarrow \underbrace{\text{erythro-dextrin} \longrightarrow \text{achroo-dextrin}}_{\text{(dextrins of decreasing complexity).}} \longrightarrow \underbrace{\text{maltose.}}_{\text{(to some extent.)}}$$

(ii) The HCl of the gastric juice may hydrolyse some sucrose.

(iii) *Amylase* of the pancreatic juice rapidly converts all forms of starch and dextrins completely into maltose. It acts in an alkaline medium, and its digestive activity is increased by the presence of the bile salts.

(iv) Succus entericus contains three classes of enzymes, *invertase* (*sucrase*), *maltase*, and *lactase*, which convert disaccharides into monosaccharides as follows :

$$\text{Sucrose} \xrightarrow{\textit{Invertase.}} \text{Glucose} + \text{Fructose.}$$

$$\text{Maltose} \xrightarrow{\textit{Maltase.}} \text{2 Glucose.}$$

$$\text{Lactose} \xrightarrow{\textit{Lactase.}} \text{Glucose} + \text{Galactose.}$$

[1] For general discussion see Soskin and Levine, *Carbohydrate Metabolism*, Chicago, 1946. Bell, *Introduction to Carbohydrate Biochemistry*, 2nd ed., London, 1948.

For an account of the general methods used in metabolic studies (including the use of isotopes), see p. 904.

[2] The reaction of saliva as secreted is faintly acid ; if the CO_2 in solution is allowed to escape by exposure to the air the reaction becomes faintly alkaline.

Pentoses are liberated as an end-product of the digestion of nucleic acid (p. 879) and from the partial digestion of pentosans.

The end-products of carbohydrate digestion are, therefore, monosaccharides, by far the most important of which is glucose.

Bacteria in the large intestine may convert some glucose into methane, CO_2, and other products.

Absorption.—(1) MECHANISM OF ABSORPTION.—Absorption of the monosaccharides from the lumen of the small intestine into the portal circulation takes place in two ways: (i) By *simple diffusion* of sugar, due to the higher concentration in the gut than in the blood.

(ii) By the *active intervention* of the living intestinal cells. The process concerned resembles that responsible for the reabsorption of glucose from the lumen of the renal tubule into the blood; it involves *phosphorylation* (p. 845), *i.e.* the union of the sugar with phosphate; it only concerns the "physiological" hexoses, *i.e.* glucose, galactose, and fructose (and not the non-physiological hexoses like mannose, or the pentoses). If the intestinal mucous membrane is killed, depressed by cooling to a low temperature, or poisoned by phloridzin (p. 32) or iodo-acetic acid, "active" absorption ceases and simple diffusion alone takes place.

(2) RATE OF ABSORPTION.—The rate of absorption by the *living* intestinal mucosa is more rapid than when diffusion alone is taking place. In man the measured maximal rate of absorption from a 50 cm. loop of jejunum per hour was: glucose, 8 g.; galactose, 9·5 g.; fructose, 5 g.[1] The corresponding rates for mannose or pentose which are absorbed solely by diffusion would be (judging from experiments on rats) about 1 g. The following factors influence the rate of absorption:

(i) The *state of the mucous membrane* ((1) (ii) *supra*) and the length of *time* during which the carbohydrate is in contact with it. Absorption is thus depressed in *diarrhœal* conditions (because of hurry), in *enteritis* and in *cœliac disease* (the nature of the mucosal disturbance in this condition is unknown). (The effects of widespread *excision* of the small intestine and of *gastro-colic fistula* are discussed on p. 797.)

(ii) *Rôle of Endocrines*.—(*a*) *Thyroid*: thyroxine acts directly on the intestinal mucosa stimulating absorption. The rate of glucose absorption is thus depressed in myxœdema accounting for the flattened glucose tolerance curve, and stimulated in hyperthyroidism giving a diabetic-looking glucose tolerance curve (cf. p. 922).

(*b*) *Anterior pituitary*: affects absorption solely through its influence on the thyroid; hyperpituitarism induces thyroid overactivity and hypopituitarism induces thyroid atrophy, with the results to be expected from (*a*) *supra*.

(*c*) *Adrenal cortex*: glucose absorption is depressed in adrenal cortex deficiency; it can be restored to normal *without the use of cortical extracts* by a *high salt diet* which raises the Na^+ level in the blood to normal. Impaired glucose absorption can thus result from an abnormal state of the intestinal cells secondary to the altered *electrolyte* content of the body fluids which is characteristic of adrenal cortex deficiency (p. 955).

(*d*) *Insulin*: has no effect on absorption.

[1] It is not known why fructose is less well favoured than glucose or galactose.

(iii) *Rôle of Vitamins.*—Various members of the vitamin-*B* complex (p. 1024), *e.g.* thiamine, pantothenic acid, and pyridoxine, promote absorption.

Fate of Hexose Sugars.—The hexoses are brought to the liver from the intestine in the portal blood. *Galactose* is converted exclusively in the liver into glucose (directly or via glycogen); fructose is converted both in the liver and in the muscles into glucose. The *glucose* which is absorbed from the bowel or is formed as a result of these isomerizations is:

(i) *Stored as glycogen* in the liver (about 100 g.) and in the skeletal muscles (about 400 g.).

(ii) "*Dissimilated*" (broken down by complete oxidation) in all the tissues to yield energy; on an average the complete oxidative breakdown of 1 g. of carbohydrate to CO_2 and H_2O yields energy equivalent to approximately 4000 cal. (=4 Cal.).

(iii) *Converted into fat* and stored as such in the fat depots (p. 874).

It should be noted that some of the intermediate products of carbohydrate breakdown may be aminated to *form amino-acids* (pp. 853, 889).

Glycogen.—Glycogen is a suitable form in which to store carbohydrate:

(i) Being insoluble it exerts no osmotic pressure and so does not disturb the intracellular fluid content.

(ii) It has a higher energy level than a corresponding weight of glucose (p. 845).

(iii) It is readily broken down under the influence of enzymes, (*a*) into glucose in the liver (to maintain the blood glucose) or (*b*) into lower intermediates in many tissues (including the liver) to yield energy.

1. **Muscle Glycogen** is formed from the circulating blood glucose; the concentration in resting muscle is 0·7–1·0%. The rate of formation of muscle glycogen is increased by a rise of blood glucose and by insulin (p. 911); it is depressed after removal of the pancreas, and is restored to normal by injection of insulin.

Muscle glycogen is consumed during muscular activity (p. 429); when muscle is rested after activity the glycogen store is built up again from blood glucose. Muscle glycogen cannot readily be converted back into blood glucose; thus even during the fatal hypoglycæmia which results from hepatectomy the muscle glycogen may not completely disappear (p. 824).

2. **Liver Glycogen.**—The presence of glycogen in the liver can be readily demonstrated as follows. An animal is anæsthetized, the liver is prepared for excision, frozen *in situ* with liquid air, removed, plunged into boiling alkali to destroy the proteins and enzymes, and then ground up. The solution contains glycogen (which can be precipitated with alcohol), but practically no glucose. On the other hand, if the liver is deprived of its blood-supply *in situ*, or kept warm for some time after removal from the dead animal it contains no glycogen, but instead large amounts of glucose are found. Clearly, glycogen is normally stored in the liver and is readily converted under ischæmic or anærobic conditions into glucose. The average liver glycogen concentration is 5% (cf. p. 827).

(1) FORMATION OF LIVER GLYCOGEN.—Glycogen is formed in the liver and muscles via the stages indicated in Fig. 551 (p. 838).

Two of the reactions are *irreversible*:

(i) The conversion of blood glucose to glucose-6-phosphate (by reaction with adenosine triphosphate under the influence of the enzyme *hexokinase*);

this reaction " introduces the glucose into the metabolic machinery of the cell " (p. 844).

(ii) The conversion of glucose-6-phosphate to blood glucose ; this occurs *in the liver only* and by the action of an entirely different enzyme system to that in (i), namely *phosphatase* (an enzyme which hydrolyzes hexose-phosphates to hexose and inorganic phosphate (cf. p. 845).

(2) SOURCES OF LIVER GLYCOGEN.—Liver glycogen is formed from :

(i) The hexose monosaccharide end-products of carbohydrate digestion *i.e.* glucose, fructose, galactose.

(ii) " Intermediates " of carbohydrate breakdown, especially lactic acid and pyruvic acid (p. 852).

(iii) Glycerol derived from the hydrolysis of neutral fat (p. 874).

(iv) Any " intermediate " derived from the breakdown of the amino-acids (of proteins) which enter into the so-called " metabolic pool " (p. 888). There is no doubt that in this way glycogen (and glucose) may be formed from many amino-acids after they have been deaminated.

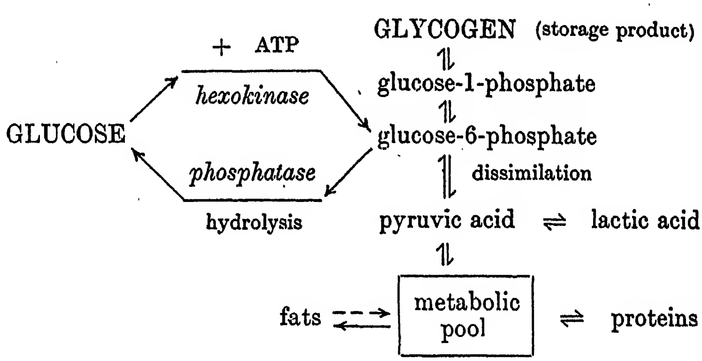


FIG. 551.—Blood Glucose and the Formation and Breakdown of Liver Glycogen.

The quickest way of building up liver glycogen, however, is by raising the blood glucose level ; this is more rapidly and effectively done by intravenous injection of glucose than by oral administration.

(3) RÔLE OF LIVER GLYCOGEN.—(i) First and foremost it is the only immediately available *reserve of blood glucose.*[1]

(ii) A high liver glycogen level *protects* the liver cell against the harmful effects of many poisons, *e.g.* carbon tetrachloride, ethyl alcohol, arsenic, and various bacterial toxins (p. 828).

(iii) Certain forms of *detoxication* (p. 828) are directly influenced by the liver glycogen level :

(*a*) detoxication by *glucuronic acid* of substances which contain OH— groups, including the sterols (*e.g.* pregnanediol (p. 1081)) ;

[1] The word " sugar " is commonly used in this book as synonymous with " glucose," *e.g.* in such expressions as " blood sugar " or " sugar in urine." Glucose is sometimes referred to as dextrose. Occasionally " sugar " is used to refer to dietary monosaccharide or disaccharide, *e.g.* sucrose (saccharose). In the plural form, " sugars " is sometimes equivalent to carbohydrates. The meaning is made clear by the context.

(*b*) detoxication by means of *acetylation* (substitution of an acetyl, *i.e.* (CH_3CO—) group), *e.g.* acetylation of *p*-amino–benzoic acid or of sulphanilamide ; the acetyl group may be formed thus : glycogen→pyruvic acid→ " 2C unit " derived from acetate (pp. 850, 873).

(iv) The *rate of deamination* of amino-acids in the liver is depressed as the glycogen level rises ; this means that the amino-acids are preserved longer in that form and so remain available for protein synthesis in the tissues.

(v) Similarly a high level of liver glycogen depresses the *rate of ketone formation* from long chain fatty acids and thus abolishes or prevents ketosis (p. 875).

The *hormones* controlling the metabolic activities of the liver, and the key rôle of the liver in the *regulation of the blood glucose level* are considered on pp. 856–859.

Other Forms in which Carbohydrate is found or Used in the Body.—(i) *Lactose:* secreted by the lactating mammary gland in the milk ; presumably it is synthesized from blood glucose (p. 1096).

(ii) *Fructose:* found as the free sugar in fœtal blood and in seminal fluid.

(iii) *Glucuronic Acid* (as *glucuronides*)*:* formed in the liver for the " detoxication " of many alcohols, phenols, and acids (p. 828).

(iv) *Pentoses* (*ribose* and *deoxyribose*)*:* constituent parts of *nucleic acids* (p. 878) and *nucleotides*. Ribose is a constituent part of the nucleotides concerned with :

(*a*) *hydrogen transport, e.g.* coenzymes-I and -II (di- and tri-phosphopyridine nucleotides respectively) (p. 841, Fig. 552) ; flavin-adenine dinucleotide from flavoproteins (p. 854) ;

(*b*) *phosphorylation, e.g.* adenosine mono-, di-, and tri-phosphates (AMP, ADP, and ATP respectively) (p. 842, and Fig. 553).

(v) *Muco-polysaccharides:* found in vitreous humour, synovial fluid, and viscous secretions generally, and in cartilage, tendon, and bone. These complex polysaccharides contain *glucuronic acid* and *hexosamines* (*glucosamine* or *galactosamine* (*chondrosamine*)) as part of their molecular structure. *Heparin* (p. 144) and *chondroitin* (the cementing material of connective tissue) are muco-polysaccharides which also contain *sulphate* groups ; *hyaluronic acid* (a highly viscous substance which can be extracted from vitreous humor, umbilical cord, and skin) contains *no* sulphate. Many muco-polysaccharides are also combined more or less firmly with specific *proteins* to form the *muco-* or *glyco-proteins, e.g. mucins* from gastric mucosa, *anterior pituitary hormones* (p. 930), and some *serum proteins*. Specific muco-polysaccharides appear to be responsible for certain antigenic reactions.

(vi) *Galactolipides* (*cerebrosides*)*:* found in brain and nervous tissue. On hydrolysis they yield *galactose*, a *fatty acid*, and a complex nitrogenous base called *sphingosine* ; their function is unknown.

DISSIMILATION OF CARBOHYDRATES IN TISSUE CELLS

The useful term *dissimilation* has been introduced to describe all the breakdown processes undergone by a foodstuff through many intermediate stages until the final end-products are reached ; in the case of the carbo-

hydrates the end-products are CO_2 and H_2O. The opposite transformation, *i.e.* the building-up of storage and structural materials from dissimilation products, is termed *assimilation*.

Before discussing the problem of carbohydrate dissimilation in detail, some chemical considerations must be briefly reviewed.

Enzyme Systems.[1]—All the metabolic reactions to be described depend upon the catalytic activity of *enzyme systems*. These consist of a soluble protein (*enzyme-protein*, *apoenzyme*) together with various accessory substances. If the latter is a simple ion (*e.g.* Mg^{++}, PO_4''') it is called a *cofactor* and usually accelerates the enzyme action; if it is a complex *organic*, but *non-protein*, substance it is called a *coenzyme* and usually acts as an essential intermediate carrier of products of the enzyme-catalysed reaction. The enzyme-protein is generally specific for a particular chemical reaction or type of reaction; a particular coenzyme, however, may act as a carrier in a number of different systems (*i.e.* it is less specific). Certain proteins (the *conjugated* proteins) consist of a protein moiety chemically combined with an organic, but non-protein, unit called the *prosthetic group* of the protein. With enzyme-proteins of this type, the prosthetic group acts as an intermediate carrier of products of the reaction, *i.e.* the enzyme has a "built-in" coenzyme; examples are the flavoproteins and cytochrome oxidase (p. 854).

Biological Oxidations.—It is characteristic of many forms of tissue oxidation that the process does *not* involve the addition of atmospheric oxygen to the substrate, but the *addition of water* ($+H_2O$) to form a "hydrate" followed by *dehydrogenation* (removal of hydrogen) ($-2H$), leaving an atom of O attached to the substrate. For a given reaction, a specific enzyme (a *dehydrogenase*) catalyses the release of 2H from the "hydrate" in the presence of a suitable *hydrogen acceptor*, *i.e.* a coenzyme (Co) which unites readily with the hydrogen. Examples are the oxidation of phosphoglyceraldehyde to phosphoglyceric acid (p. 843) and the oxidative decarboxylation of pyruvic acid (p. 850), both of which are of the general type:

$$\mathrm{R{-}CHO} \xrightarrow{+H_2O} \underset{\text{“ hydrate ”}}{\mathrm{R{-}CH(OH)_2}} \xrightarrow{\text{enzyme}+\text{Co}} \mathrm{R{-}COOH} + \mathrm{Co.2H}$$

(where R = an organic grouping)

The addition of *phosphoric acid* rather than of water may occur to give phosphate derivatives. Many of the reactions are *reversible*.

The commonest coenzyme acceptor for dehydrogenase reactions is a complex derivative of nicotinamide (one of the vitamins of the *B* group, p. 1026) called *coenzyme-I* (Co-I; also called cozymase, or diphosphopyridine nucleotide, DPN). Its constitution is shown in Fig. 552; it is believed that in the reduced form the hydrogen is attached to the nicotinamide moiety. There is also a corresponding triphosphopyridine nucleotide (TPN, coenzyme-II). These are instances of a "micro-dietary" component (nicotinamide) being indispensable to the body (*i.e.* being a *vitamin*) because it participates in a fundamental metabolic process and yet cannot be synthesised.

The reduced-coenzyme is ultimately re-oxidised by passing its hydrogen along a *chain of hydrogen carriers*, finally to form water by reaction with molecular oxygen (p. 854).

[1] Dixon, *Multi-enzyme Systems*, Cambridge, 1949. Sumner and Myrbäck, *The Enzymes*, N.Y., 1950, 1951.

Energy Transfer.—All chemical reactions involve energy exchanges. The energy values are customarily expressed in *heat* units (calories [1]); a chemical reaction which liberates heat is said to be *exothermic*, and one which takes in heat is *endothermic*. However, the energy which can be made available for the performance of useful work by a chemical reaction is not always equivalent to the heat liberated, irrespective of the efficiency of the process. It is the *available* (*metabolically-usable*) *energy* which is of importance in metabolic processes. Reactions releasing usable energy are termed *exergonic* and those requiring external energy to be supplied are termed *endergonic*. Most dissimilation (catabolic) reactions are exergonic. Thus the complete oxidation of 1 g.-mol. (180 g.) of glucose to CO_2 and H_2O liberates energy equivalent to 686,000 cal. of heat; in other words, the energy content of glucose is 686,000 cal./g.-mol. above the energy content of its oxidation

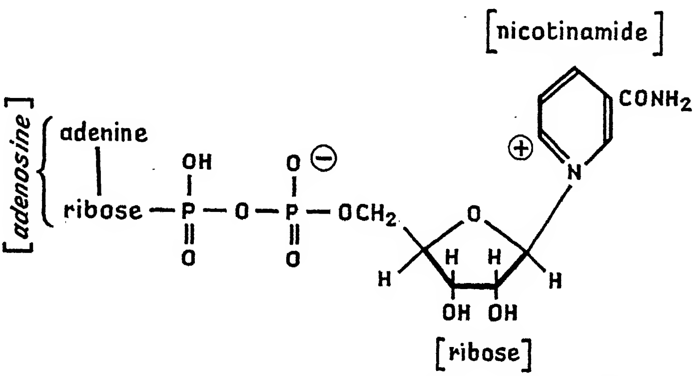


FIG. 552.—Coenzyme-I. (Diphosphopyridine nucleotide, DPN)

products. To reverse the reaction and synthesise 1 g.-mol. of glucose from CO_2 and H_2O (as occurs in plants during photosynthesis), energy equivalent to 686,000 cal. must be supplied to the system from an outside source.

A complex molecule has a higher energy content than the atoms or simpler molecules from which it is built because of the energy of formation of the chemical bonds which hold it together; this *bond-energy* is liberated when the bonds are broken. The bond-energy of the link between two given atoms of a molecule naturally depends on the structure of the *whole* molecule. This point is well illustrated by reference to the very important group of organic phosphates, which have the general formula $R.PO(OH)_2$, where R represents an organic radical derived, for example, from glucose, creatine, etc. (see p. 842).

When most organic phosphates are hydrolysed (liberating phosphoric acid), energy equivalent to about 2000–3000 cal./g.-mol. is made available as heat. In these cases, the bond broken is called a *low-energy phosphate bond*; the formulae of phosphates with such bonds are written R—ph, where —ph represents the phosphate group joined to the rest of the molecule

[1] One g.-calorie (cal.)=heat to raise the temperature of 1 g. of water through 1°. One kg.-calorie (Cal.)=1000 cal. Cf. footnote on p. 375.

by a low-energy bond.[1] But when some organic phosphates (of special structural types) are hydrolysed, energy equivalent to about 12,000 cal./g.-mol. is made available. These phosphates contain *high-energy phosphate bonds* ; their formulae are written R∼ph, where ∼ph represents a phosphate group joined to the rest of the molecule by a high-energy bond.

The metabolically-important phosphates of each type are :

(i) *Low-energy :* glucose–1–phosphate and glucose-6-phosphate ; fructose-diphosphate ; phosphoglyceraldehyde ; monophospho-glyceric acid ; adenosine monophosphate (AMP).

(ii) *High-energy :* the second and third phosphate bonds of adenosine triphosphate (ATP) ; creatine phosphate ; diphospho-glyceric acid ; phospho-pyruvic acid ; acetyl phosphate.

Thus :

phosphoric acid HO—PO(OH)$_2$	HO—ph	1 low-energy phosphate bond
creatine phosphate	creatine∼ph (full structure Fig. 570, p. 892)	1 high-energy phosphate bond
adenosine monophosphate (AMP)	adenosine—ph	1 low-energy phosphate bond
adenosine diphosphate (ADP)	adenosine—ph∼ph	1 low-energy, 1 high-energy phosphate bond
adenosine triphosphate (ATP)	adenosine—ph∼ph∼ph (full structure Fig. 553)	1 low-energy, 2 high-energy phosphate bonds

RÔLE OF HIGH-ENERGY PHOSPHATES.—These high-energy phosphates are important in metabolism for the following reasons :

(*a*) The phosphate group may be *transferred* directly to another organic molecule without much of the high energy of its bonding being dissipated as heat. The product is a *phosphorylated* molecule ; it may or may not have a high-energy phosphate bond, depending on its structure, but it has a total energy content exceeding that of the *non*-phosphorylated molecule by

[1] The symbol ph for phosphate is not to be confused with *p*H (log of hydrogen-ion concentration (p. 89)).

3000–12,000 cal./g.-mol. Examples are the interaction of ATP with glucose (p. 845)

glucose + adenosine—ph$\sim$ph$\sim$ph (ATP)
$\longrightarrow$ glucose—ph (glucose-6-phosphate) + adenosine—ph$\sim$ph (ADP)

and with creatine (pp. 429, 893)

creatine + adenosine—ph$\sim$ph$\sim$ph
$\rightleftharpoons$ creatine$\sim$ph (creatine phosphate) + adenosine—ph$\sim$ph (ADP)

It must again be emphasized that a molecule formed by phosphate transference from a compound with a high-energy phosphate group does *not necessarily have a high-energy phosphate group itself*, though the whole molecule is raised to a higher energy level; only the special phosphate structures listed previously have high-energy phosphate bonds.

(*b*) If the overall dissimilation processes of metabolism are imitated in a test-tube, the liberated energy is dissipated as heat. In the body, this energy of dissimilation, instead of immediately being lost as heat, is used to *synthesise compounds containing high-energy phosphate bonds* (*e.g.* ATP); these compounds are stored, and the energy "locked-up" in them is utilised as required (see (*c*) below). The mechanism of the process is believed to be as follows: phosphate compounds containing *low-energy* phosphate bonds undergo reactions (usually oxidations) that convert them into other derivatives containing *high-energy* phosphate bonds; the resulting phosphate groups, with their attendant high energy, are then used to phosphorylate ADP, giving ATP. The energy of the dissimilation has thus been "channelled" into the products of the reaction. For example, the direct oxidation *in vitro* of phosphoglyceraldehyde to phosphoglyceric acid liberates energy as heat; but in the body, this stage of the dissimilation of carbohydrate proceeds thus (see also p. 848):

ph—G.CHO (phosphoglyceraldehyde) $\xrightarrow{+\ \text{HO—ph (phosphoric acid)}}$ ph—G.CH(OH)(O—ph)

$\downarrow$ *oxidation* (—2H)

ph—G.CO.O$\sim$ph (diphospho-glyceric acid) $\xrightarrow{+\ \text{ADP}}$ ph—G.COOH (phospho-glyceric acid) + ATP (usable energy as $\sim$ph)

[ADP takes up the $\sim$ph to form ATP] [G = —OCH_2.CHOH—]

(*c*) High-energy phosphate appears to be the *sole source of energy* that cells can use *directly*. It is used: (α) to effect chemical *synthesis*; (β) to perform *work* (muscular, osmotic, secretory); (γ) to liberate *heat* (by

hydrolysis) (maintaining the body temperature but thereby dissipating the metabolic energy). The energy of muscle contraction is believed to be derived directly from ATP (p. 429).

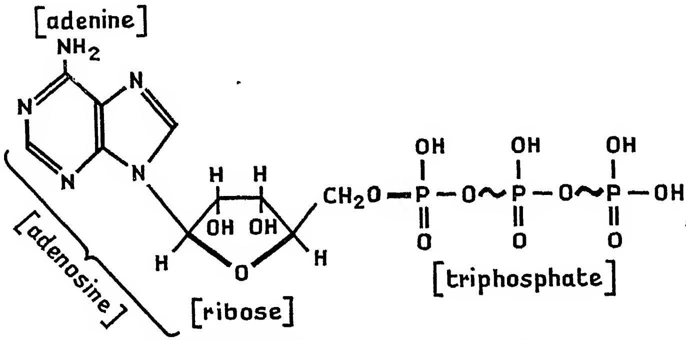


FIG. 553.—Adenosine Triphosphate (ATP).

Structure of Glucose, Fructose, and Glycogen.—The structural formula of glucose is shown in Fig. 554. It contains a *pyranose* ring system (5 C atoms and 1 O atom); the 6 C atoms of glucose, numbered from (1) to (6) as shown, carry hydroxyl (—OH) substituents and can form phosphate esters (—O—ph).

Fructose (an isomer of glucose) also contains a pyranose ring, but its phosphate esters are derived from a modified, more reactive, *furanose* ring system (4 C atoms and 1 O atom).

Glycogen consists of many hundreds of glucose units linked together (with elimination of the elements of water) through *glucoside* linkages (C—O—C). The linkage may be between $C_{(1)}$ of one unit and $C_{(4)}$ of an adjacent unit (a 1 : 4 linkage) giving a straight chain, or between $C_{(1)}$ of one unit and $C_{(6)}$ of another (a 1 : 6 linkage) giving a branched arrangement.

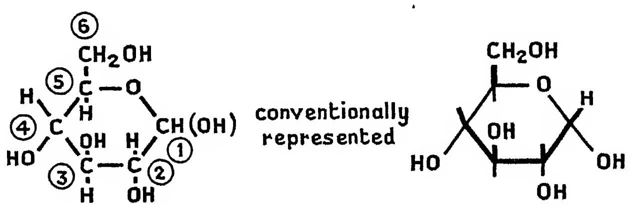


FIG. 554.—The Structure of Glucose.

Both types of linkage (and possibly others) occur in glycogen, which is thus a large ramified molecule (Fig. 555).

Uptake of Blood Glucose to form Glucose-6-Phosphate. Hexokinase Reaction.—(1) Glucose is not directly metabolized in the body. It is first raised to a higher energy level (and thus "taken into the metabolic machinery" of tissue cells) by conversion into glucose-6-phosphate. This

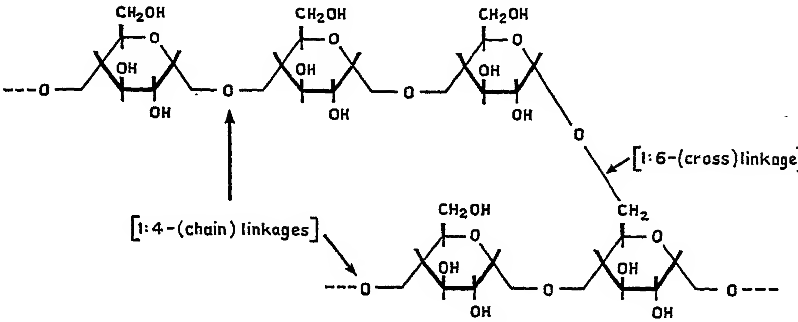


FIG. 555.—The Structure of Glycogen.

uptake of phosphate is called *phosphorylation*,[1] not to be confused with phosphorolysis (see p. 847).

The responsible enzyme is called *hexokinase*; it "specifically activates"[2] the glucose molecule to receive phosphate from adenosine triphosphate (ATP), thus:

$$\text{glucose} + \underset{\text{(ATP)}}{\text{adenosine—ph}\sim\text{ph}\sim\text{ph}}$$

$$\xrightarrow{\textit{hexokinase}} \underset{\substack{\text{(glucose-6-}\\ \text{phosphate)}}}{\text{glucose—ph}} + \underset{\text{(ADP)}}{\text{adenosine—ph}\sim\text{ph}}$$

Glucose-6-phosphate has a low-energy phosphate bond; but since the whole molecule is at a higher energy level than glucose, the hexokinase reaction is *irreversible*. Free blood glucose is regenerated (mainly in the liver) from glucose-6-phosphate by direct *hydrolysis*, under the influence of an enzyme, *phosphatase*; this reaction also is *irreversible*.

$$\text{glucose-6-phosphate} + \text{water}$$

$$\xrightarrow{\textit{phosphatase}} \text{glucose} + \text{phosphate (+ energy wasted as heat)}$$

It has been suggested that the effects of anterior pituitary factors, adrenal corticoids and insulin on glucose metabolism may be due in part to their influence on the hexokinase reaction (pp. 913, 937).

(2) If ATP is to continue to function as an *energy donor* in the hexokinase reaction, it must be steadily re-formed by the phosphorylation of ADP. In muscle, the necessary high-energy phosphate groupings (~ph) are provided by the stored creatine phosphate (pp. 429, 893), thus:

$$\underset{\text{(creatine phosphate)}}{\text{creatine}\sim\text{ph}} + \underset{\text{(ADP)}}{\text{adenosine—ph}\sim\text{ph}}$$

$$\rightleftharpoons \text{creatine} + \underset{\text{(ATP)}}{\text{adenosine—ph}\sim\text{ph}\sim\text{ph}}$$

(3) "Below" the level of glucose-6-phosphate, the metabolic histories of glucose and glycogen are identical. The energy levels of glucose-6-phosphate,

[1] Cf. phosphorylation in the absorption of glucose from the intestine (p. 836) and the renal tubule.

[2] This resounding term, "activates," gives the impression of conveying meaning without of course explaining what is as yet not understood, *i.e.* *how* the enzyme really acts.

follow apply mainly to processes in skeletal muscle and liver, but probably similar changes occur in other tissues.

The dissimilation occurs in two main stages :

(A.) The breakdown of glycogen (or of glucose via glucose-6-phosphate) to pyruvic acid.

(B.) The further transformations of pyruvic acid.

Fig. 556 shows the known steps of stage A.

(A.) Glycogen or Glucose to Pyruvic Acid.—(1) GLYCOGEN→GLUCOSE-1-PHOSPHATE (process : *phosphorolysis* ; enzymes : *phosphorylases* ; reaction is reversible). Phosphorolysis involves splitting the glucoside linkages of glycogen by orthophosphoric acid (H_3PO_4, H—O—ph), attaching —O—ph at $C_{(1)}$ of one unit, and H— to the adjacent unit (Fig. 557). Phosphorolysis

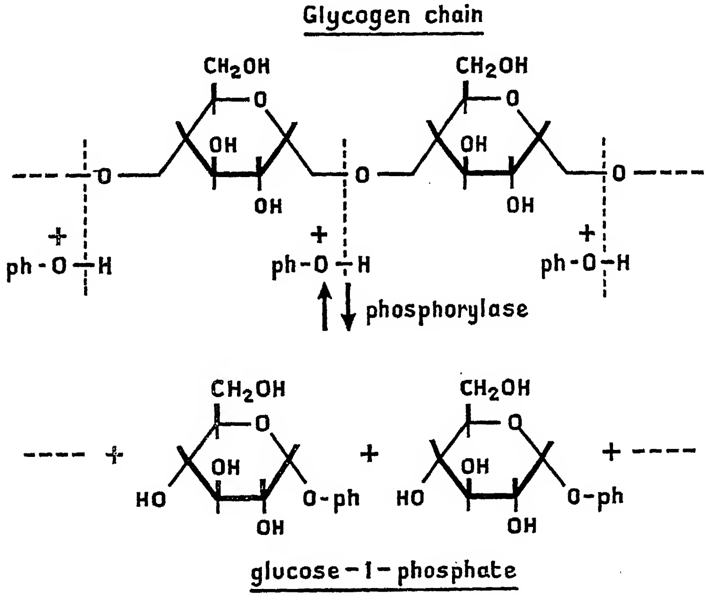


FIG. 557.—The Phosphorolysis of Glycogen to give Glucose-1-Phosphate.

should be compared with hydrolysis, where splitting is accompanied by the addition of the elements of *water*, H—O–H (*e.g.* the hydrolysis of starch or glycogen by amylases). Phosphorolysis *differs* from hydrolysis in the following ways : (i) hydrolysis liberates energy as waste heat ; phosphorolysis gives a product (in this case, glucose-1-phosphate) with approximately the *same energy level* as the substrate (glycogen) ; (ii) hydrolysis is usually irreversible because of the energy change ; phosphorolysis is *reversible*.

(2) GLUCOSE-1-PHOSPHATE→GLUCOSE-6-PHOSPHATE (process : *intramolecular phosphate transfer* ; enzyme : *phosphogluco-mutase* ; reaction is reversible). The low-energy phosphate group is transferred from $C_{(1)}$ to $C_{(6)}$. The metabolic paths of glycogen and glucose join here (see p. 845).

(3) GLUCOSE-6-PHOSPHATE→FRUCTOSE-6-PHOSPHATE (process: *isomerization*; reaction is reversible); its significance is unknown.

(4) FRUCTOSE-6-PHOSPHATE→FRUCTOSE-1 : 6-DIPHOSPHATE (process: *phosphorylation*), raising the molecule to a higher energy level, utilising a high-energy phosphate bond (∼ph) from ATP.

fructose—ph + adenosine—ph∼ph∼ph
⟶ ph—fructose—ph + adenosine—ph∼ph

(5) FRUCTOSE-DIPHOSPHATE→2 Mols. of TRIOSE-PHOSPHATE (process: *cleavage*; enzyme: *zymohexase*; reaction is reversible, little energy change). The triose (3 C) compound *phospho-glyceraldehyde* is ultimately formed.

(6) PHOSPHO-GLYCERALDEHYDE→PHOSPHO-GLYCERIC ACID. This reaction can be considered as occurring in three stages (hypothetically-distinct, see section on phosphate energy, p. 843):

(*a*) phosphoglyceraldehyde + H_3PO_4 → diphosphoglyceraldehyde.
(*b*) diphosphoglyceraldehyde + Coenzyme-I (Co-I)
⟶ diphosphoglyceric acid + Co-I.2H
(*c*) diphosphoglyceric acid + adenosine diphosphate (ADP) →
phosphoglyceric acid + adenosine triphosphate (ATP).

Stage (*a*) involves the non-enzymic *addition of phosphoric acid* to phosphoglyceraldehyde. The pathways for glucose and *glycerol* (from fats) join here.

In stage (*b*), *dehydrogenation* of the addition product occurs (enzyme: *triose-phosphate dehydrogenase*) to give diphosphoglyceric acid, which has one of its phosphate groups associated with a *high-energy bond*. The hydrogen from the dehydrogenation is passed directly to *coenzyme-I* (Co-I) (see Fig. 556, p. 846). It might be thought that the sequence of reactions described above would have to stop when all the coenzyme-I (Co-I) had been reduced to Co-I.2H. However, it is permitted to continue because the reduced coenzyme is *reoxidised* by transferring its 2H either (α) directly to *pyruvic acid* (under anærobic conditions only, see below), or (β) through a chain of *H acceptors* until it is finally disposed of by uniting with molecular oxygen to form water (for details, see p. 854).

In stage (*c*) the high-energy phosphate group of diphosphoglyceric acid (together with its energy bond) is *transferred* to a molecule of ADP, raising it to the higher energy level of ATP. The energy liberated in passing from phosphoglyceraldehyde to phosphoglyceric acid has thus been retained within the system and used to synthesise ATP from ADP and phosphate.

(7) PHOSPHO-GLYCERIC ACID→PHOSPHO-PYRUVIC ACID (processes: *rearrangement*, followed by *dehydration*). Phosphopyruvic acid has a *high-energy phosphate bond*; phosphoglyceric acid has not.

(8) PHOSPHO-PYRUVIC ACID→PYRUVIC ACID (process: *transference of high-energy phosphate* from phosphopyruvic acid to ADP, giving ATP).

It should be remembered that each glucose unit of glycogen gives rise to *two* molecules of triose and subsequent products. Thus the dissimilation of one glucose unit of glycogen to two molecules of pyruvic acid generates energy sufficient to be bound as *four* high-energy phosphate bonds in four newly-formed ATP molecules (two formed at step 6(*c*) and two at step 8). But one ∼ph from ATP is used (in step 4) to raise the energy content of the reactant to a level high enough to initiate the chain of dissimilation reactions. Therefore the *net* release of metabolically-available energy is *three high-energy*

phosphate bonds (approx. 36,000 cal.) per g.-mol. of glucose. If the process starts from free blood glucose, the *net* release is only *two high-energy phosphate bonds*, since the ~ph of one ATP is used in the hexokinase reaction (p. 844).

(B.) Fate of Pyruvic Acid.—Pyruvic acid, $CH_3.CO.COOH$, is a key substance in metabolism. It is a metabolic stage which is reached by *all the foodstuffs* (*i.e. protein and fat, as well as carbohydrate*) *either directly or through common intermediates* (see Figs. 558, 562, 569).

(1) Conversion of Pyruvic Acid to Lactic Acid.—This conversion is of importance because it occurs in skeletal muscle working under conditions of absolute or relative *oxygen lack, e.g.* in isolated muscle contracting in an atmosphere of nitrogen, or in working muscles during maximal exercise. In the *presence* of adequate oxygen supplies, *no lactic acid is formed*; pyruvic acid is broken down through complicated intermediate steps to CO_2 and H_2O (p. 850), and the reduced coenzyme (Co-I.2H) is steadily reoxidised to Co-I so that it can continue to function as hydrogen-carrier in stage A. (6b). In the *absence* of adequate oxygen supplies all the Co-I would soon be put out of action as it would all have been converted to the reduced form, Co-I.2H. But in these circumstances, pyruvic acid acts as a *temporary H carrier* (acting in place of the molecular oxygen which is not available), *i.e.* pyruvic acid *dehydrogenates* (oxidizes) the (reduced) Co-I.2H back to (oxidized) Co-I; the latter can then continue to act as hydrogen carrier in the energy-mobilizing reactions leading to the formation of more pyruvic acid. The pyruvic acid itself is reduced to *lactic acid.*

$$\text{Co-I.2H} + \underset{\text{pyruvic acid.}}{CH_3.CO.COOH} \rightarrow \text{Co-I} + \underset{\text{lactic acid.}}{CH_3.CHOH.COOH}$$

Thus *under anaerobic conditions*, all the muscle glycogen can be broken down into *lactic acid* to yield energy without the Co-I being completely saturated with H and so put out of action (Fig. 556).

The reaction pyruvic acid⇌lactic acid is *reversible:* the same enzyme and coenzyme are involved whichever way the reaction is moving. If a muscle forming lactic acid is re-supplied with enough oxygen, or the lactic acid circulates to a point of greater oxygen availability, the reaction is reversed and lactic acid is reconverted to pyruvic acid, thus :

$$CH_3.CHOH.COOH + \text{Co-I} \longrightarrow CH_3.CO.COOH + \text{Co-I.2H}$$
$$\text{Co-I.2H} \longrightarrow \text{Co-I} + 2H$$
$$2H \text{ (through many intermediate carriers, see p. 854)} + \text{mol. } O \rightarrow H_2O$$

The enzyme concerned is called *muscle lactic dehydrogenase* simply because it has been mainly studied in the lactic acid→pyruvic acid direction. When the enzyme is catalysing the reverse reaction it is catalysing the *addition* of H and would be more appropriately called a *hydrogenase* (but enzyme terminology must be taken as it is, with all its imperfections).

It is important to realize that the details of glycogen dissimilation down to pyruvic acid set out in A (1)–(8) above come mainly from experiments done under *anærobic* conditions, *i.e.* conditions under which only lactic acid, and *not* pyruvic acid, appears. On the whole, opinion favours the view that the stages in the ærobic pathway leading to pyruvic acid are the same as the initial stages leading to lactic acid, *i.e.* that the scheme is followed oxidatively as well as anærobically, but there is no certainty about this ; alternative pathways from glycogen to pyruvic acid may play some part.

(2) DISSIMILATION OF PYRUVIC ACID TO CO_2 AND H_2O.—*Mechanism of CO_2 Formation. Decarboxylation.*[1]—It has been emphasized that molecular oxygen does not unite directly with hydrogen of the substrate, but that the hydrogen must pursue a complicated route to reach molecular oxygen to form water (p. 854). Likewise, *molecular oxygen never unites directly with carbon of the substrate to form carbon dioxide.* CO_2 is always liberated by a process of enzymic *decarboxylation, i.e.* CO_2 is split off from the carboxyl group (—COOH) of intermediate organic acids.

The details of these reactions are obscure, but the following points appear established:

Decarboxylation may be either non-oxidative or oxidative:

(i) *Non-oxidative, i.e.* CO_2 alone is split off, thus:

$$\underset{\text{pyruvic acid.}}{CH_3.CO.COOH} \longrightarrow \underset{\text{acetaldehyde}}{CH_3.CHO} + CO_2$$

(ii) *Oxidative, i.e.* in addition to the liberation of CO_2, *dehydrogenation* (oxidation) also occurs, thus:

$$\underset{\text{pyruvic acid.}}{CH_3.CO.COOH} + H_2O + \underset{\text{(coenzyme)}}{Co} \longrightarrow \underset{\text{acetic acid.}}{CH_3.COOH} + CO_2 + Co.2H$$

Oxidative decarboxylation is irreversible; it generates energy which is bound as the high-energy phosphate bond of ATP, and requires the presence of a hydrogen carrier (usually Co-I). A further coenzyme, *cocarboxylase*, is required for *both* types of decarboxylation; cocarboxylase is a phosphate derivative of *aneurin* (*thiamine, vitamin-B_1*)—another example of a vitamin participating in a fundamental enzyme system (p. 1026).

The first dissimilation reaction of pyruvic acid is its oxidative decarboxylation to acetic acid and CO_2; complex changes, including two further decarboxylations, follow, and ultimately all the C of the substrate is liberated as CO_2 (Fig. 558).

CO_2 Assimilation.—Some non-oxidative decarboxylations are *reversible, i.e.* the enzymes involved can catalyse CO_2 uptake (*assimilation*) as well as CO_2 liberation. Thus the following important reaction of pyruvic acid is known to occur (enzyme: *β.carboxylase*), giving *oxaloacetic acid*:

$$\underset{\text{pyruvic acid.}}{CH_3.CO.COOH} + CO_2 \rightleftharpoons \underset{\text{oxaloacetic acid.}}{COOH.CH_2.CO.COOH}$$

CO_2 is not always a final waste product of metabolism; it may be used as a molecular building material, as in this and many other examples (cf. urea, p. 887; purines, p. 897).

Details of Pyruvic Acid Dissimilation. Tricarboxylic Acid Cycle.—One might have thought that pyruvic acid would simply lose successive C atoms in the form of CO_2, as set out above. But the normal route of dissimilation turns out to be a quite unexpected and complicated one. The route is as follows (see Fig. 558):

(i) By a process of *CO_2 assimilation* (*carboxylase reaction*) one molecule of pyruvic acid takes up one molecule of CO_2 to form *oxaloacetic acid* (4C; two COOH groups = a dicarboxylic acid).

(ii) Another molecule of pyruvic acid undergoes *oxidative decarboxylation* to *acetic acid* (2C; one COOH group).

[1] Ochoa, *Physiol. Rev.*, 1950, *31*, 56.

(iii) The acetic acid then unites with the oxaloacetic acid to form a 6C compound with three COOH groups, *i.e.* a *tricarboxylic acid,* almost certainly *citric acid.*

(iv) By means of two decarboxylations and related oxidations, this 6C acid is degraded through a 5C compound to a 4C compound.

The *equivalent* of the three C atoms present in one of the original molecules of pyruvic acid is thus converted into, and given off as, three molecules of CO_2.

(v) The 4C residue (from the 6C compound) is oxidized back into oxaloacetic acid ; the latter (as in (iii)) unites with a *second* molecule of acetic acid (from another molecule of pyruvic acid) to re-form the 6C acid, and so on. The whole process is thus *cyclical,* with oxaloacetic acid acting " catalytically." (The above steps are indicated in Fig. 558 by the appropriate numbers.) The system, which has been demonstrated in many tissues, is variously labelled the *tricarboxylic acid cycle,* the *citric acid cycle,* or the *Krebs cycle.*[1]

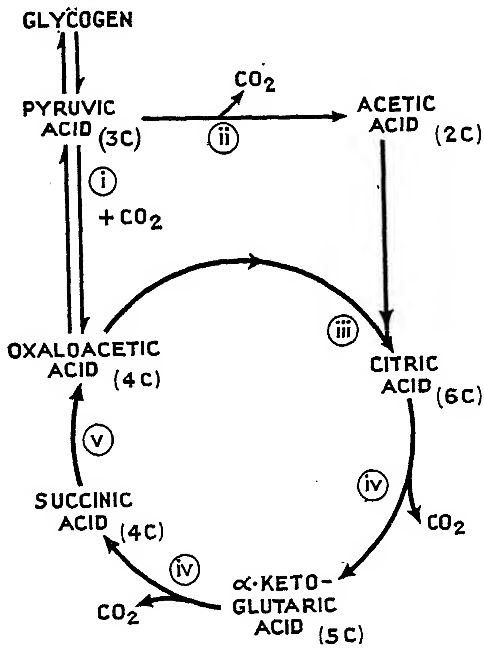


FIG. 558.—The Krebs Citric Acid Cycle (cf. Figs. 559, 569).
(Diagram by J. B. Jepson.)

The structures of some of the known intermediates of the cycle are shown in Fig. 559. The citric acid cycle is, properly speaking, a *system of enzymes,* not of substrates ; however, it is found convenient to think of it in terms of

[1] After H. A. Krebs, who originally propounded, and found experimental evidence for, the system. This investigator also gives his name to another cyclical enzyme system—the *Krebs urea cycle* (p. 886).

substrates, since most of the individual enzymes concerned have not been isolated.

Each time the circuit is completed, one molecule of pyruvic acid ($CH_3COCOOH$) is dissimilated, 3 molecules of CO_2 are evolved and 10 atoms of H are passed to oxygen (5 O) via coenzyme hydrogen acceptors; the cycle only acts under *ærobic* conditions.

During this process of dissimilation the total energy of oxidation of the pyruvic acid is made available. *The energy so released is bound as the high-energy phosphate bonds of newly formed ATP:* the mechanisms are presumably similar to those operating during the breakdown of glucose to pyruvic acid (p. 848). *It is found that the complete dissimilation of* 2 *molecules of pyruvic*

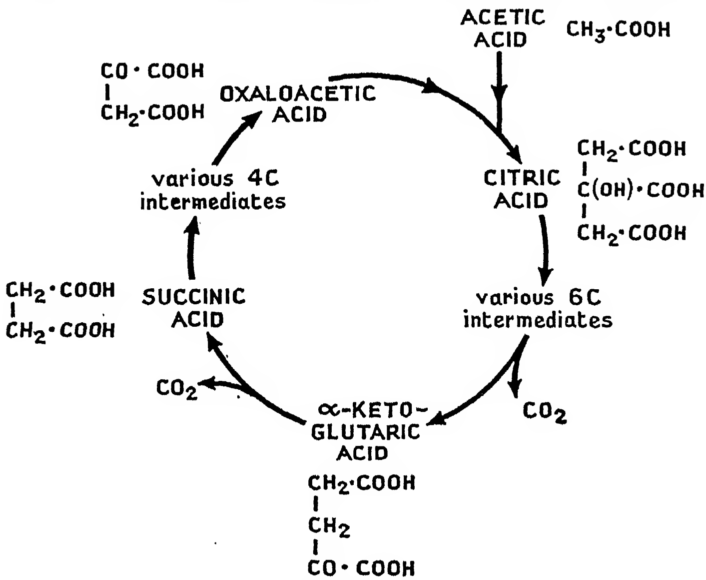


FIG. 559.—Some important Intermediates of the Citric Acid Cycle. Cf. with Fig. 569. (Diagram by J. B. Jepson.)

acid (i.e. 1 *glucose unit*) *results in the formation of approximately* 40 *high-energy phosphate bonds* (40 *molecules of ATP*), *or more than* 10 *times the energy yield of the breakdown from glucose to pyruvic acid.* Thus the ærobic phase yields 10 times more energy in useful form than does the anærobic phase.

The cycle is probably unidirectional as shown; components of the cycle can "escape" as pyruvic acid by reversal of the carboxylase reaction, step (i), Fig. 558 (see (4) below).

(3) RECONVERSION OF PYRUVIC ACID TO GLUCOSE OR GLYCOGEN.—The dissimilation of glycogen (or glucose-6-phosphate) down to pyruvic acid is *reversible.*[1] Thus pyruvic acid (or any earlier intermediate) can be *rebuilt* into glycogen or blood glucose (*glucogenesis*). The dissimilation process from glucose to pyruvic acid liberates energy; therefore the reverse process

[1] This is not to say that all the individual enzyme reactions are directly reversible—*most* of them are, and "by-passes" are available for the others.

of glucogenesis (which occurs mainly in the liver) will require an equal amount of energy to be supplied back. This energy is provided by ATP; this substance which contains high energy phosphate bonds has itself been built up from other energy-liberating reactions, *e.g.* from the *complete oxidation* in the liver of a small proportion of the available pyruvic acid. The complete oxidation of 2 molecules of pyruvic acid or lactic acid provides enough energy to rebuild 10 molecules of pyruvic acid into 5 molecules of glucose.

(4) UTILIZATION OF PYRUVIC ACID IN PROTEIN AND FAT SYNTHESIS.—(i) It will be shown (p. 888) that many *amino-acids* after deamination are transformed directly or indirectly into acids which participate in the Krebs cycle (Fig. 558). In this way they "go round the cycle" to form oxaloacetic acid and are either dissimilated or else give pyruvic acid, and so may ultimately be built into glycogen or glucose.

(ii) Conversely, these constituent acids of the cycle (formed from carbohydrate via pyruvic acid) may be *aminated* (gain NH_2) to form amino-acids which may then be used in *protein* synthesis, thus:

$$\begin{array}{c}\text{pyruvic}\\\text{acid}\end{array} \rightleftharpoons \underset{\substack{\text{oxaloacetic}\\\text{acid}}}{\begin{array}{l}CH_2.COOH\\|\\CO.COOH\end{array}} \underset{\text{deamination}}{\overset{\text{amination}}{\rightleftharpoons}} \underset{\substack{\text{aspartic}\\\text{acid}}}{\begin{array}{l}CH_2.COOH\\|\\CH(NH_2)COOH\end{array}} \rightleftharpoons \text{proteins}$$

(iii) Similarly, carbohydrate and *fat* metabolism have a common meeting point in the Krebs cycle (Fig. 562). Pyruvic acid is oxidatively decarboxylated to acetic acid (p. 850); acetic acid (or some closely related 2C compound) is also the chief product from fatty acid oxidation (p. 871), bringing fat into the carbohydrate pathway. The reverse reaction will synthesize fatty acid by combination of acetic acid units obtained from pyruvic acid (p. 874).

(iv) The *glycerol* for combination with fatty acids to give neutral fats can also be synthesised from carbohydrate by the reduction and hydrolysis of phosphoglyceraldehyde, an intermediate on the chain of reactions from glucose to pyruvic acid (p. 848).

In these ways, proteins, fats, and carbohydrates are interconvertible.

Common Metabolic Pool.—The elucidation of the metabolic pathways of the three main groups of foodstuffs and body components has led to the following important conception: that, after initial modifications in the first stages of metabolism, the carbohydrates, proteins, and fats *become incorporated into a system of common carbon fragments no longer distinguishable as to origin.* This system is termed the "*common metabolic pool,*" and is based mainly on the citric acid cycle. Various reversible and irreversible processes, forming other cycles, govern the operation of the central metabolic system. If one type of fragment is lacking when needed, or is present in excess, then appropriate cycles are set in motion in such a manner as will produce that fragment, or dissimilate it.

The entry and exit of carbohydrates, fats, and proteins to and from the "common metabolic pool" are discussed above and on pp. 874 and 887.

Similarly, each distinct reversible metabolic cycle constitutes a "body pool" for a given compound, *e.g.* the body glucose pool, the amino-acid pool; the use of isotopically labelled molecules has contributed largely to our knowledge of these metabolic cycles (p. 908). The size of an individual pool can be measured by an isotopic dilution technique. The amount of

substance thus measured is not solely the amount of free substance distributed throughout the body but also includes the amount which could be manufactured reversibly by the operation of other cycles. For example, the "body glucose pool" would include some portion of the liver glycogen as well as all the free glucose of the body fluids. The meaning of the word "pool" is thus determined by the measurement techniques that are employed.

Transmigration of Hydrogen to Unite with Molecular Oxygen to form Water.—When we took temporary leave of the hydrogen released from various substrates by the action of dehydrogenases (pp. 848, 852) it had become attached to a coenzyme hydrogen acceptor, namely coenzyme-I (DPN) (or in certain cases, coenzyme-II (TPN) or a flavin (see below)). The subsequent transfer of hydrogen to molecular oxygen involves the mediation of a *flavoprotein*, *cytochrome*, and *cytochrome oxidase*, a group of substances shown by the spectroscope to be widely present in cells.

FLAVOPROTEINS.—These are enzymes which catalyse the *transfer of hydrogen* from reduced-coenzymes to the *prosthetic group* of the enzyme (p. 840). This prosthetic group is a yellow pigment called *flavin* (or more exactly, flavin adenine dinucleotide, FADN); it is a derivative of *riboflavin* (vitamin-B_2)—yet another example of the metabolic function of *B*-group vitamins (p. 1028).

$$\text{Co.2H} + \text{flavin} \rightarrow \text{Co} + \text{flavin.2H}$$

CYTOCHROME.—This pigment[1] is a conjugated protein carrying an *iron-containing porphyrin* prosthetic group called cytochrome porphyrin. The latter is allied to, but not identical with, the hæm group of hæmoglobin (p. 173) and it is bound to a protein which is not globin. In view of the fact that hæm and the porphyrin group of cytochrome resemble each other so closely, it is remarkable that the hæmo-proteins derived from them differ so fundamentally in their functions. Cytochrome (Cyto.) accepts hydrogen from reduced flavin, thus being itself reduced (for mechanism see below):

$$\text{flavin.2H} + \text{Cyto.} \rightarrow \text{flavin} + \text{Cyto.2H}$$

The hydrogen is then transferred to the enzyme *cytochrome oxidase*, thus regenerating oxidized cytochrome.

CYTOCHROME OXIDASE.—This enzyme is so called because it catalyses the transfer of hydrogen from reduced cytochrome to its own "built-in" carrier (prosthetic group), which is another *iron-porphyrin*, similar to, but not identical with, hæm or cytochrome porphyrin. Reduced cytochrome oxidase transfers its hydrogen to molecular oxygen to form water (for mechanism, see below). In this way cytochrome oxidase is restored to its oxidized form (CyOx), and the hydrogen from the substrate reaches its final resting-place by combining with oxygen derived from the air to form *water*.

$$\begin{array}{lllll} \text{Cyto.2H} + \text{CyOx} & \longrightarrow & \text{Cyto.} + \text{CyOx.2H} \\ \text{CyOx.2H} + \tfrac{1}{2}\,O_2 & \longrightarrow & \text{CyOx} + H_2O \end{array}$$

HYDROGEN TRANSFER IN TERMS OF ELECTRON TRANSFER.—The above scheme of hydrogen transfer is an oversimplified account. In terms of modern chemical theory, the fundamental process involved is *electron transfer*.

[1] Several related cytochromes appear to exist, with similar functions but different specificities. We shall treat them as one.

A hydrogen atom removed during a dehydrogenation reaction undergoes "*acidic ionisation*" to give a *proton* (positively charged, denoted by H^+) and an *electron* (equally but oppositely charged, denoted by ε').

$$\underset{\text{atom}}{H} \rightleftharpoons \underset{\text{proton}}{H^+} + \underset{\text{electron}}{\varepsilon'}$$

The term *oxidation* is applied to a reaction in which a substance *loses electrons* (as well as, of course, to a reaction in which O is gained or H lost); similarly, a *gain of electrons* is termed *reduction.* Thus, considering the two forms of combined iron, ferrous (Fe^{++}, reduced form) and ferric (Fe^{+++}, oxidized form), we have

$$\text{(reduced)}\quad Fe^{++} \underset{+\varepsilon'}{\overset{-\varepsilon'}{\rightleftarrows}} Fe^{+++}\quad \text{(oxidized)}$$

This is equally the case with iron-containing cytochrome (Cyto-Fe^{+++}) and cytochrome oxidase (CyOx-Fe^{+++}):

$$H \text{ (from substrate via Co and flavin)} \longrightarrow H^+ + \varepsilon'$$

$$\text{Cyto-}Fe^{+++} + \varepsilon' \longrightarrow \text{Cyto-}Fe^{++}$$

$$\text{Cyto-}Fe^{++} + \text{CyOx-}Fe^{+++} \longrightarrow \text{Cyto-}Fe^{+++} + \text{CyOx-}Fe^{++}$$

Reduced cytochrome oxidase ,CyOx-Fe^{++}, has the unique property of transferring an electron to molecular oxygen:

$$2\ \text{CyOx-}Fe^{++} + \tfrac{1}{2}\,O_2 \longrightarrow 2\ \text{CyOx-}Fe^{+++} + O''$$

Protons from the initial ionization then react with the negatively-charged oxygen to give water:

$$2\,H^+ + O'' \longrightarrow H_2O$$

REGULATION OF BLOOD GLUCOSE. LIVER AND CARBOHYDRATE METABOLISM.

Normal Blood Glucose.—The normal morning fasting level of blood glucose is 80–100 mg–%. After a meal containing carbohydrate the level rises temporarily to a varying extent, but does not normally exceed 180 mg–%. The changes in blood glucose which follow the ingestion of 50 g. of glucose (*glucose tolerance curve*) are described on p. 922.

As glucose is a readily diffusible substance it is distributed fairly uniformly throughout the 50 L of body fluids, both extracellular and intracellular. With a blood glucose level of 100 mg–%, the total glucose content of the whole circulating blood (5 L) is only 5 g., but the total glucose content of the body fluids is 50 g. Isotope dilution methods show that the body "glucose pool" also includes the amount of glycogen which is available for rapid and reversible conversion into blood glucose. The comparatively large size of the "glucose pool" enables it to serve as a "buffer" to minimize variations in the blood glucose level. For example: irrespective of the carbohydrate intake, the peak rate of sugar absorption from the intestine in man is not more than 120 g. per hour. If all this 120 g. of sugar were absorbed *instantaneously*, the rise in the sugar content of the blood and body fluids would only be 120 g./50 L=240 mg–%. Similarly the "glucose pool" can function as a "reservoir" when glucose is being withdrawn by the tissues, *e.g.* during activity.

Glucose normally *enters* the blood (and body fluids) from the intestine and

the liver; it *passes out* into the liver and tissues where it is deposited as glycogen, dissimilated, or transformed into fat. Many and complex mechanisms are in constant operation to preserve the relative constancy of the blood glucose level under all normal circumstances. They are briefly reviewed on p. 860.

The rôle of the liver is discussed below.

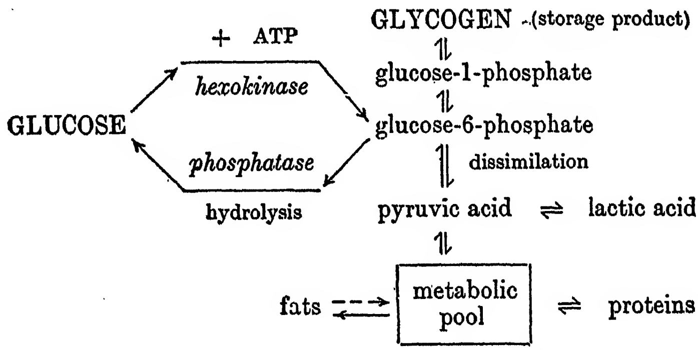
Relation of Liver to Regulation of Blood Glucose.—Glucostatic[1] Function of Liver.—The liver in the intact animal maintains the normal level of the blood glucose in three main ways:

(i) By regulating the reversible reaction blood glucose ⇌ liver glycogen. It should be noted that glycogen can also be formed from other hexose sugars (fructose and galactose) and from products of partial dissimilation of glycogen, especially lactic acid (p. 837).

(ii) By regulating new glucose formation (*neoglucogenesis*) mainly from non-nitrogenous residues of amino-acids derived from protein, and also to some extent from the glycerol derived from fat.

(iii) By regulating the removal of glucose from the blood and its deposition as fat in the liver. The fat can subsequently be stored as fat in the depots, or transformed into ketones and distributed to the tissues for dissimilation.

Methods (i) and (ii) are discussed more fully below. The reactions concerned are summarized thus:



(1) Regulation of the Reversible Reaction, Glucose ⇌ Glycogen.—(i) Glycogen deposition (*glycogenesis*) normally occurs whenever the blood sugar level rises, *e.g.* when hexose sugars are being absorbed from the intestine after a carbohydrate meal.

(*a*) The liver is affected *directly* by the *concentration* of blood sugar reaching it: a local rise of blood sugar increases glycogen deposition; on the other hand a fall of local blood sugar stimulates glycogen breakdown and thus sets up the "secretion" of glucose into the blood (*infra*).

(*b*) Glycogenesis (at a constant blood glucose level) is stimulated by *insulin* (and *adrenal corticoids*) (p. 946). In insulin deficiency (diabetes mellitus) the liver glycogen content is low; it is restored to normal by insulin treatment.

(ii) Glycogen is converted into blood glucose (*glycogenolysis*) whenever

[1] "Glucostatic": the word is used here to mean "maintaining the normal range of blood sugar level."

the blood glucose tends to fall. The fall of blood glucose not only acts *directly* on the liver but also on the central nervous system, and so *indirectly* on the liver. A fall of blood glucose stimulates brain centres, leading to increased *sympathetic* activity and additional secretion of *adrenaline* and, possibly, to increased secretion of diabetogenic factor of the *anterior pituitary*. The sympathetic impulses to the liver acting in conjunction with the hormones, stimulate glycogenolysis ; in this way, the restoration of the blood glucose level is aided.

Liver glycogen content is diminished, as might be expected, in muscular exercise, exposure to cold, starvation, and following the hypoglycæmia which is produced by insulin injection in the *normal* animal. The liver response in each case is a compensatory reaction to the hypoglycæmia. Stimulation of the sympathetic supply to the liver (p. 711), injection of adrenaline (p. 727), thyroxine (p. 979), diabetogenic anterior pituitary extracts (p. 937), all of which stimulate glycogenolysis, lead to a *rise* of blood glucose in the otherwise *normal* animal.

(2) Regulation of Neoglucogenesis.—By *neoglucogenesis* is meant new glucose formation, *i.e.* formation of glucose from non-carbohydrate sources ; the sources are mainly protein (non-nitrogenous residues of certain amino-acids) and glycerol derived from fat.

(i) The *occurrence* and *hepatic site* of neoglucogenesis can easily be proved.

(*a*) In the fasting animal, the blood glucose is kept within normal limits almost until death occurs. As the liver glycogen store is exhausted in about 24 hours, the blood glucose utilized by the tissues must be replaced from non-carbohydrate reserves, such as tissue protein, or reserve fat, or both.

(*b*) Removal of the liver leads rapidly to a fatal hypoglycæmia (p. 824). This is not due simply to the " amputation " of the liver glycogen because, as just stated, exhaustion of the liver glycogen is not followed by hypoglycæmia in the fasting animal with an *intact* liver. In the fasting animal then, the blood glucose is being *manufactured by the liver* from non-carbohydrate sources ; the liver is the only site of neoglucogenesis.

(ii) The apparent *scale* of neoglucogenesis has been calculated as follows : in the *liverless* dog, glucose must be infused at the rate of 0·25 g. per kg. per hour to maintain the blood glucose at the normal level. This is then the rate at which glucose (when it is the only metabolite supplied) is used by the tissues of the *liverless* animal. If it could be assumed that glucose is the only metabolite supplied to the tissues by the liver in the starving *normal* animal, then this figure would also represent the rate at which the liver is manufacturing new glucose in the starving animal. For a man the corresponding rate would be 420 g. of *new glucose daily*. But there is no doubt that the liver in starvation is also forming and supplying *ketone bodies* to the tissues on a considerable scale ; the amount of neoglucogenesis would be correspondingly smaller.

(iii) *Sources of Neoglucogenesis.*—Neoglucogenesis takes place mainly from *protein* and to a smaller extent from the *glycerol from fat.*

(*a*) *Protein.*—Evidence is set out on p. 888 that many of the amino-acids can be converted into glucose. It is reasonable to suppose that such a transformation is an important source of the blood glucose in the fasting animal (and perhaps to some extent in the normal animal). An injection of amino-acids into the liverless animal does not raise the blood glucose ; the liver is thus the only site of new glucose formation from amino-acids.

(*b*) *Fat.—Glycerol,* from the hydrolysis of fats, can be transformed into blood glucose by the liver, via phosphoglyceraldehyde which lies on the reversible pathway of glucose dissimilation (p. 848). It requires the hydrolysis of 8–10 g. of depot or liver fat to provide enough glycerol to yield 1 g. of blood glucose by this route.

There has been much debate as to whether, and if so on what scale, the *fatty acids* derived from fat can also be transformed by the liver into blood glucose in the starving animal. In the absence of any known pathway for this transformation it is concluded here that the amount of such conversion is very small (p. 874).

There is no doubt, however, that fat plays an important *indirect* rôle in maintaining the blood glucose level :

(*α*) Fatty acids are partially dissimilated in the liver to give ketone bodies which are circulated to the tissues ; there they are completely dissimilated to yield energy, thus reducing the " demand " on the blood glucose.

(*β*) Fatty acids are completely dissimilated in the liver liberating energy which is used locally in the resynthesis of glucose from (for example) circulating lactic acid (p. 853).

(iv) *Control of Neoglucogenesis.*—Neoglucogenesis is regulated so as to maintain the *normal* blood glucose level ; if this level rises, new glucose formation is decreased, if it falls glucose formation is increased. Neoglucogenesis is regulated by the *blood glucose level itself,* acting in two ways :

(*a*) directly on the liver,

(*b*) indirectly by affecting the secretion of certain *hormones.*

The hormones which influence neoglucogenesis are :

(*α*) insulin, which depresses it,

(*β*) the diabetogenic factor of the anterior pituitary, certain adrenal corticoids, and thyroxine, which stimulate it (p. 859).

(3) Control of Blood Glucose by the Liver.—Claude Bernard regarded the liver as an organ of *internal* secretion, its internal secretion being glucose.[1] As was shown above, the substances which give rise to blood glucose, *i.e. glucogens,* are carbohydrate (glycogen), protein and glycerol of fat ; the last two might be labelled *neo-glucogens.*

When the blood glucose level tends to *fall,* the reaction

$$\underset{\text{(glycogen, amino-acid residues and glycerol).}}{\text{All Glucogens}} \longrightarrow \text{Glucose}$$

proceeds at a rate which just compensates for the withdrawal of glucose from the blood. When the blood glucose level tends to *rise,* the reaction (glucogens →glucose) ceases and the reactions (glucose→glycogen) and (glucose→fatty acids) commence and counteract the rise by removing the excess glucose from the blood.

An analogy can be drawn between this blood glucose regulating activity of the liver and the rôle of the hypothalamus in maintaining normal body temperature. Certain centres in the hypothalamus act as a *thermostat*: they are sensitive to minute variations in body temperature and produce appropriate reactions to keep the temperature within the normal " pre-fixed "

[1] When Claude Bernard gave to the substance which he had extracted from the liver the name glucogène he was thinking of it as a " *glucogen,*" *i.e.* a substance which *generates glucose* to be turned out into the blood stream. Strangely enough the English term glycogen which originally was equivalent to the French glucogène has come to refer principally to a form of storage.

range. The liver can be described as a *blood glucostat,* reacting to maintain the blood glucose within the range of the normal. In both instances there is a small time lag before the full effect of the reactions is felt, so that some deviation from the " normal " may occur under conditions of stress.

(4) HORMONAL CONTROL OF THE GLUCOSTATIC FUNCTION OF THE LIVER. —The regulation of the blood glucose by the hormones depends only in part (and not wholly) on their action on the liver. The hormones concerned can be divided into two groups with roughly opposing actions, namely (i) *insulin,* and (ii) the *diabetogenic group* which includes the diabetogenic factor of the anterior pituitary, adrenal corticoids, and thyroxine.

(i) *Action of Insulin on Liver.*—Insulin increases the deposition of glycogen, depresses neoglucogenesis, and increases the transformation of glucose into long chain fatty acids (lipogenesis) (p. 913). In the absence of insulin, hyperglycæmia develops which is due in part to depressed glucose utilization by the tissues. The liver, however, is also at fault in the diabetic organism, because it does not react in such a manner as would tend to restore the blood glucose to normal. The " correct " hepatic response to hyperglycæmia would be increased withdrawal of glucose from the blood and decreased formation of glucose. In the diabetic, however, the formation of glycogen and of fat from glucose is depressed ; new glucose formation continues at a rate which is equal to, or even exceeds the normal rate ; in other words unnecessary amounts of glucose continue to be secreted by the liver into the blood.

(ii) *Action of Diabetogenic Group of Hormones on Liver.*—This group increases the output of glucose by the liver from glycogen and non-carbohydrate sources (see pp. 937, 945, 979). (The individual hormones, however, differ in the details of their action on the liver.) In the absence of these hormones a fall of the blood glucose level is a less effective stimulus than normally in promoting glucogenesis. There is consequently a diminished secretion of glucose by the liver into the blood, which is inadequate to maintain the blood glucose level in the fasting animal ; *hypoglycæmia* develops.

(iii) Normally the activity of the two antagonistic groups of hormones is nicely adjusted as to preserve the normal blood glucose level. A rise of blood glucose increases insulin secretion (p. 917) and presumably depresses the secretion of diabetogenic factor ; a fall of blood glucose depresses insulin secretion and stimulates secretion of diabetogenic factor. If one or other group of hormones is put out of action the normal balance is lost, as explained above.

(iv) In the absence of both insulin and the anterior pituitary factor (its principal antagonist), a state of affairs which is found in an animal from which both the pancreas and the anterior pituitary have been removed (*Houssay animal*), the main regulators of the hepatic glucostat are not available. The glucostatic activity of the liver is then determined solely by the direct action on it of the blood glucose level. Uncontrolled by the action of the hormones, the liver reacts imperfectly and with a much greater time lag. In the fasting Houssay animal, hypoglycæmia develops (which may be rapidly fatal), while after a carbohydrate meal considerable hyperglycæmia occurs (cf. p. 938).

(v) The *nervous* (*sympathetic*) supply to the liver is of minor importance in blood sugar regulation ; denervation of the liver does not significantly impair the efficiency of the glucostatic mechanism.

Regulation of Blood Glucose.[1]—NORMAL BLOOD GLUCOSE.—See p. 855.

The principal regulatory factors are as follows:

(1) ALIMENTARY CANAL.—Ingested polysaccharides are *slowly* broken down into glucose, so that absorption is never very rapid. If glucose itself is ingested in large amounts it is held up for a considerable time in the stomach; if taken in still greater quantities it is rejected by vomiting (p. 929).

(2) LIVER.—The fundamental rôle of the liver as a blood glucostat was fully discussed on pp. 856–858.

(3) TISSUES.—The activity of the tissues may either disturb or help to restore the blood glucose level. Tissue activity withdraws glucose for dissimilation and lowers the blood glucose. On the other hand a rise of blood glucose, both by a direct action and by calling forth a secretion of insulin, increases dissimilation of glucose, deposition of glycogen, and the transformation of glucose into depot fat.

(4) RÔLE OF HORMONES.—The blood glucose level is automatically regulated by appropriate responses of certain ductless glands. Their rôle is discussed in the following places: *insulin* (p. 911); *anterior pituitary* (p. 937); *adrenal cortex* (p. 945).

(5) RÔLE OF NERVOUS SYSTEM.—The central nervous system plays an accessory rather than a major part in these processes. Removal of the cerebral cortex or of the entire cerebrum or decerebration through the mid-brain all lead to prolonged hyperglycæmia and glycosuria; this result is likewise attributed to sympathetic overaction.[2]

If any considerable change in the blood sugar occurs, symptoms are produced; those of hypoglycæmia are described on pp. 824, 915; those of hyperglycæmia are very uncertain.[3]

CAUSES OF HYPERGLYCÆMIA.—Hyperglycæmia is due to more glucose entering the blood than leaving it. This can occur under the following circumstances:

(i) *Excessive glucogenesis in the liver* from glycogen, amino-acids or glycerol of fat. It may, therefore, follow injection of adrenaline (p. 727), anterior pituitary extracts (p. 937) or thyroxine (p. 979); or it may develop in asphyxia or during anaesthesia. Hyperglycæmia may occur in emotional stress, in exophthalmic goitre (p. 991), in acromegaly (p. 935), or in the experimental conditions known as thyroid and hypophyseal diabetes (p. 918). For the same reason, lesions of the brain in man may be associated with hyperglycæmia (see above).

(ii) *Depressed utilization of glucose.* After pancreatectomy, injury to the islet by alloxan, administration of thyroid (metathyroid diabetes) or anterior pituitary extracts (metahypophyseal diabetes) (p. 918), or in clinical disorders of the pancreas, the hyperglycæmia is due partly to diminished glucose utilization and storage (but partly also to excessive glucogenesis).

At certain levels of hyperglycæmia glycosuria develops (p. 927).

CAUSES OF HYPOGLYCÆMIA.—Hypoglycæmia can be produced by: (i) excess insulin (p. 914); (ii) hepatectomy or diminished liver activity (p. 824).

[1] Soskin, *Physiol. Rev.*, 1941, *21*, 140.

[2] The classical observations of Claude Bernard showed that puncture of the floor of the fourth ventricle causes hyperglycæmia; his results were probably due to stimulation of central pathways which control the sympathetic supply to the liver and the adrenal medulla.

[3] The effects of hyperglycæmia may be mainly secondary to the changes it produces in renal function, *e.g.* excess loss of water and ions in the urine with consequent anhydræmia and circulatory failure.

FAT METABOLISM[1]

Chemistry of Fats.—The word *lipide* is widely used to refer to certain chemical compounds which are insoluble in water but are soluble in the so-called fat solvents (*e.g.* alcohol, ether, benzene, chloroform).[2] The lipides of physiological interest can be classified as follows:

(1) SIMPLE LIPIDES: the most important are the *neutral fats* (*triglycerides*), *i.e.* glyceryl esters of fatty acids.

Triglycerides are formed from one molecule of glycerol and *three* molecules of fatty acid, thus:

$$\begin{array}{lcl} CH_2OH & HOOC.R & CH_2O.OC.R \\ | & & | \\ CHOH \;+ & HOOC.R' \longrightarrow & CHO.OC.R' \\ | & & | \\ CH_2OH & HOOC.R'' & CH_2O.OC.R'' \\ \text{Glycerol.} & \text{Fatty acids.} & \text{Triglyceride.} \end{array}$$

where R, R′, and R″ represent three (possibly different) radicals of fatty acids. *Diglycerides* and *monoglycerides* are formed from one molecule of glycerol and *two* molecules or *one* molecule of fatty acid respectively. The most common of the many known fatty acids are:

(i) *Palmitic acid*, $CH_3(CH_2)_{14}COOH$ (a 16 C acid).

(ii) *Stearic acid*, $CH_3(CH_2)_{16}COOH$ (an 18 C acid).

(iii) *Oleic acid*, $CH_3(CH_2)_7CH{=}CH(CH_2)_7COOH$ (an 18 C *unsaturated* acid).

(i) and (ii) are fully saturated acids; (iii) is an *unsaturated* acid, containing one double bond, —CH=CH—. Practically all the natural fatty acids contain an *even* number of C atoms (as might be expected from their mode of synthesis, p. 873).

Waxes are esters of fatty acids with *long-chain alcohols* (instead of with glycerol, as in neutral fats).

(2) COMPOUND LIPIDES: complex compounds formed from fatty acids, glycerol (or related substances), and various *nitrogen*-containing bases, and often containing *phosphate* groups. They are integral parts of the general cell structure (p. 866); they are present in large amounts in nervous tissue (p. 482); they are employed in fat transport (pp. 864, 868).

The chief types of compound lipide are:

(i) *Phospholipides* (*phosphatides*), containing glycerol, 2 molecules of fatty acid (generally unsaturated), phosphate, and a nitrogen-containing base (*choline*, $(CH_3)_3\overset{+}{N}CH_2CH_2OH$, in the case of the *lecithins*; *ethanolamine* (*cholamine*), $NH_2CH_2CH_2OH$, in the case of the *cephalins*).

(ii) *Sphingomyelins*, containing fatty acid, phosphate, choline, and a complex base (*sphingosine*), but *no* glycerol.

[1] Bloor, *Biochemistry of Fatty Acids and Lipids*, New York, 1943. Burr and Barnes, *Physiol. Rev.*, 1943, *23*, 256. Symposium on "Triglycerides in Human Nutrition," *Brit. J. Nutrition*, 1949, *3*, 350.

[2] In this book, however, the common British practice is followed which is to use the word "fat" more vaguely and less correctly as more or less synonymous with lipide (see *e.g.* blood fat, p. 865; tissue fat, p. 866). The terms *lipin* or *lipoid* are sometimes used instead of lipide.

(iii) *Galactolipides* (*cerebrosides*), containing the monosaccharide galactose, fatty acid, and sphingosine, but *no* phosphate or glycerol.

(3) ASSOCIATED LIPIDES: These are of two main types: (i) Those components (the *split fats*) that are obtained by the *hydrolysis* of lipides (*i.e.* glycerol, fatty acids, soaps). *Soaps* are salts of fatty acids, and are obtained by the hydrolysis of fats in alkali (*saponification*).

(ii) Those components that are associated with the lipides in tissue extracts simply because they are dissolved by the fat solvents: mainly (*a*) the *steroids* (cf. p. 1074), *e.g.* hormones of the ovary, testis, adrenal cortex; cholesterol (esters of cholesterol with fatty acids are called *cholesterides*); (*b*) the *fat-soluble vitamins* (*i.e.* vitamins soluble in fats and in fat solvents) (p. 1019).

The fat of food consists mainly of neutral fat, together with small amounts of free fatty acid, lecithin, and cholesterol esters. The nutritional importance of dietary fat is discussed on p. 1055.

Digestion of Neutral Fat.—Some hydrolysis of neutral fat takes place during cooking and the process may continue to a small extent in the stomach.

STOMACH.—In exceptional circumstances, significant fat digestion may occur in the stomach. A fat-splitting enzyme (*gastric lipase*) is present in the pure gastric juice from a Pavlov pouch; in addition, pancreatic lipase may regurgitate into the stomach from the duodenum. The activity of lipase in the stomach is obviously restricted, because the enzyme is sensitive to free acid and is destroyed by exposure to 0·02% HCl for 15 minutes. Fat hydrolysis (lipolysis) may take place in the stomach in cases of *achylia gastrica* and in young *suckling* animals which ingest large quantities of milk: the fat of milk is present in an emulsified and therefore readily digested form, and it also inhibits the secretion of gastric acid.

SMALL INTESTINE.—The main digestion of fat undoubtedly occurs in the small intestine, the active enzyme being predominantly *pancreatic lipase* (which is assisted in many ways by the *bile*); *succus entericus* also has some lipolytic action.

It used to be thought that triglycerides had to be completely hydrolysed in the intestine to glycerol and three molecules of fatty acid before absorption. This is not now believed to be the case.

Lipase hydrolyses triglyceride in stages, splitting off successively each of the three attached fatty acid molecules: (i) removal of one fatty acid leaves a *diglyceride* (containing two fatty acids); (ii) removal of the second fatty acid leaves a *monoglyceride* (containing one fatty acid); (iii) removal of the third fatty acid leaves *glycerol*. Under normal conditions in the intestine, fat digestion is incomplete; the end-products are undigested triglyceride, partially split products (di-glyceride and mono-glyceride), fatty acid (derived from completely and partially split fat), and glycerol. Lipolysis is initially rapid, but soon slows down and finally stops before 30% of the fatty acids of the ingested triglyceride has been liberated.

Lipase is activated non-specifically by the *bile salts* (p. 797); it is rapidly destroyed *in vitro* by trypsin but is presumably protected from attack in the intestine. As the optimum *p*H for lipolysis is on the alkaline side of neutrality the activity of lipase is slightly hampered by the normal slight acidity of the intestinal contents (*p*H 5·5–6·5). Because of this acid intestinal *p*H, none of the fatty acid liberated by digestion is converted into soap.

At body temperature food fat is in the liquid form as an oil. The speed of lipolysis depends on the size of the *surface area* of the oil that is exposed to the enzyme. Lipolysis is thus greatly accelerated by *emulsification*; in an emulsion the fat is divided into small particles which are dispersed through the watery phase. The most satisfactory emulsion is one which is stable and in which the fat droplets are very minute. In the intestine, neutral fat is emulsified by the products of its initial digestion, *i.e.* by a mixture of the lower glycerides (di-glycerides and mono-glycerides) and "complexes" formed between the fatty acids and the bile acids (*vide infra*). These emulsions are very fine (the fat droplets are less than $0{\cdot}5\mu$ in diameter), and are stable at the *p*H and under the other conditions of the small intestine.

Fatty acids are *insoluble* in water and *indiffusible* through animal membranes. They combine, however, with the bile acids to form a loose chemical compound called a *fatty acid-bile acid complex* which is stable in the intestinal environment, water-soluble and diffusible, and can pass through the intestinal epithelium. (The bile acids are thus called *hydrotropic* substances; by uniting with an insoluble substance they make it water-soluble; probably cholesterol and the fat-soluble vitamins (-*A*, -*D*, -*E*) are also acted upon by bile acids in this way.)

Absorption of Fat.[1]—The end-products of fat digestion are thus:

(i) Mainly (70%) unhydrolyzed, water-insoluble, neutral fat which under normal conditions in the intestine has been reduced to a finely emulsified state.

(ii) Water-soluble and diffusible fatty acid-bile acid complexes.

(iii) Mono-glycerides and di-glycerides, some soluble and some emulsified.

(iv) Glycerol, which is water-soluble and diffusible.

According to Frazer, the water-soluble products, *i.e.* the fatty acid-bile acid complexes and the glycerol are *absorbed directly into the portal blood*. On the other hand the emulsified but water-insoluble neutral fat is absorbed via the lining epithelial cells of the intestine into the *lacteals* and thence into the thoracic duct and into the venous blood. This theory is called the "partition hypothesis" because it supposes that there are two distinct routes for fat absorption. Some of the experimental evidence may be cited:

(i) Fat particles, so-called *chylomicrons*, can be detected in the systemic blood under dark-ground illumination as points of light and counted; during starvation they are scanty, but after a fat meal they increase in number. If rats are given well-emulsified olive oil (a neutral fat) the lacteals appear milky, and the chylomicron count in the systemic blood rises but that in the portal blood remains low. It is concluded that the absorption of neutral fat occurs via the lacteals into the systemic circulation.

(ii) On the other hand, if a mixture of fatty acids and glycerol is administered the lacteals remain relatively free from fat but the chylomicron count in the portal blood rises sharply; the fatty acid, on reaching the liver, is retained there and the systemic chylomicron count remains low.

(iii) Only about two-thirds of the absorbed fat can be recovered by cannulating the lymphatics, it is argued that the missing third is absorbed by some other route, possibly directly into the blood stream.

(iv) If cetyl sulphate is added to neutral fat introduced into the intestine it has two main effects: it produces a very fine emulsion and it *inhibits the action of lipase*. Though in these circumstances the neutral fat cannot be

[1] Frazer, *Brit. med. J.*, 1947, ii, 641. Rapor, *ibid.*, 1949, ii, 719.

hydrolyzed to fatty acid, fat is absorbed quite rapidly; the lacteals soon become milky, and fat quickly appears in the systemic blood, whence it is deposited in the fat depots but not in the liver.

The most important prerequisite for the direct absorption of neutral fat is its thorough emulsification into particles which are less than $0{\cdot}5\mu$ in diameter; it is normally only in the presence of mono-glyceride, fatty acid, and bile salt that this fine emulsion can be maintained.

In some unknown way these emulsified neutral fat particles can penetrate into the intestinal cells. During their passage through the intestinal wall they may be temporarily transformed into *phospholipide* (*phosphatide*) before being reconverted into neutral fat to pass out of the cell into the lacteals. Phosphatide formation involves the temporary uptake of inorganic phosphate and may be catalysed by the enzyme phosphatase which is present in high concentration in the intestinal mucosa. Some of the evidence supporting the phosphatide hypothesis may be quoted:

(i) The addition of phosphate or of glycerophosphate to the intestinal contents increases the rate of absorption of fatty acids from a mixture of fatty acids, bile acids, and glycerol.

(ii) Iodoacetic acid and phloridzin, which inhibit phosphatase, depress fatty acid absorption.

(iii) The use of staining methods (of doubtful validity) purports to show that when these poisons are used fatty acid absorption is blocked because phosphorylation does not occur.

(iv) Experiments with radio-active phosphorus $^{32}P^{*}$ (as phosphate) show a high turnover of phosphate during fat absorption.

It is possible that fat absorption is influenced by the local concentration of electrolytes, by hormones and perhaps by vitamins. The adrenal cortex, by regulating the electrolytic balance in the body fluids, may be a subsidiary regulating factor (p. 955).

It has been shown that very finely emulsified liquid paraffin in which the particle size is less than $0{\cdot}5\mu$ can be absorbed to the extent of 40% by the rat intestine. Fortunately medicinal liquid paraffin, when used in cooking or as a laxative, is not sufficiently finely dispersed to be absorbed to a significant degree in man.

To summarize: efficient fat absorption requires both *lipase* and *bile salts.*

(i) *Lipase* produces sufficient lipolysis to form mono-glycerides and di-glycerides which together with the bile acid-fatty acid complex, emulsify the undigested tri-glyceride and so enable it to be absorbed.

(ii) The *bile salts* form a water-soluble complex with fatty acids thus enabling them to be absorbed; and the complex also helps (by promoting emulsification) the absorption of undigested fat as stated in (i) *supra.*

The fundamental importance of bile salts in fat absorption must be emphasized. If bile is excluded from the intestine, fat digestion proceeds fairly normally, but large amounts of fat are lost in the fæces (p. 798).[1]

[1] In the *rabbit* the pancreatic duct enters the small intestine some 30 cm. below the bile duct; fat absorption only occurs beyond the point of entry of the pancreatic duct, *i.e.* where preliminary lipolysis (liberating fatty acid and lower glycerides) has taken place. Conversely the bile duct has been ligated and the gall bladder anastomozed with the distal part of the small intestine; some fat digestion occurs proximally (owing to the activity of the lipase of the pancreatic juice), but no absorption occurs until contact with the bile has been established.

Movements of the Villi.—Several hours after a meal, the villi show a variety of movements, the most important being a shortening (owing to contraction of the contained smooth muscle fibres) which leads to compression of the central lacteal and evacuation of its fatty contents into the deeper lymphatics. The villi are stimulated mechanically, by certain constituents of the food and by an alleged blood-borne hormone labelled *villikinin*.

The lymph laden with fat is further propelled: by intestinal movements which compress the valved deeper lymphatics; by the raised intra-abdominal pressure during inspiration; and lastly, by the negative pressure in the thorax. It passes up the thoracic duct into the great veins at the root of the neck and thus into the systemic circulation.

Blood Fat.—The blood contains fat in the following forms (the figures in parentheses give the normal plasma levels in mg. per 100 c.c.):

(i) mainly as neutral fat and fatty acids (200–450);

(ii) as lecithin and other choline-containing phospholipides (150–250);

(iii) as free cholesterol and cholesterol esters (cholesterides) (total 150–250 of which 70% is the ester form).

An increase in blood-fat is called a *lipæmia*. The increase in the blood concentration of neutral fat which follows a meal is irregular and unpredictable. The blood-fat is also increased in the following conditions: (i) *starvation* (p. 900); (ii) *diabetes mellitus* (p. 924); (iii) *pregnancy*; (iv) certain types of renal disease (*nephrosis* (p. 114)); (v) *myxœdema* (p. 984).

Fate of Fat after Absorption.—After absorption, fat is treated in various ways:

(1) It undergoes *complete oxidation* in the tissues to yield energy, CO_2, and H_2O (p. 872). When 1 g. of mixed fat is completely oxidized to CO_2 and H_2O, 9·4 Cal. of heat are produced; during the metabolic dissimilation of fat, a large proportion of this amount of energy is made available to the body as high-energy phosphate bonds. "Acetic acid units" derived from the breakdown of fat can be used in *acetylation* reactions and for the *synthesis* of acetoacetic acid and certain body components (p. 873).

(2) It is *stored* (as neutral fat) in the fat depots. In contrast to the small carbohydrate reserves (0·5 kg.), fat may be stored in the body in very large amounts. On an average, fat forms over 10% of the body weight (*i.e.* about 7 kg. in an adult, equivalent to an energy reserve of 1000 Cal. per kg. of body weight, or more than a month's total food energy); in people who are over-weight the fat reserves are much bigger.

It must be emphasized that neutral fat is *not* deposited interstitially in the matrix between cells or fibres; it is taken up by the cells of *adipose tissue*, in which the cytoplasm gradually diminishes in amount as the fat accumulates, until the cell becomes a thin, cytoplasmic nucleated envelope enclosing a large fat droplet. Adipose tissue should be regarded as a highly specific tissue, *taking up* fat differentially when fat is present in excess of the body's immediate metabolic needs, and *releasing* it when required. Adipose tissue takes up fat in the same selective manner as, for example, the thyroid takes up iodine. Nothing is known about the way in which the finely emulsified fat particles in the blood enter the adipose tissue cells or, what is even more difficult to understand, how the seemingly unemulsified fat in the adipose tissue cells is discharged into the circulation. There is evidence that the activity of adipose

tissue is under hormonal control but the details of the regulation are still extremely obscure.[1]

(3) Fat is built into the *structure* of all tissues. The structural lipide consists of the following groups of lipides (see p. 861 for their chemical composition):

(i) *Lecithins* (and the related *cephalins*).

(ii) *Cholesterides* (cholesterol esters of fatty acids).

The constituent fatty acids in these two groups are mainly unsaturated (unlike the fatty acids in depot fat, many of which are fully saturated). Lecithin and cholesterides are essential constituents of all cell membranes; lecithin is a component of the medullary sheath of nerve fibres.

(iii) Certain specialized lipides like the *sphingomyelins* and *cerebrosides* of the central nervous system.

(iv) The specialized *steroid hormones* of the ovary, testis, and adrenal cortex (pp. 1074 ff., 959).

Structural lipides are as integral a part of the cell architecture as are proteins; they constitute the *élément constant* of the total body lipide (in contrast to the *élément variable* or neutral fat of the fat depots). In starvation, the neutral fat in the depots is called upon and used for producing energy, but structural lipides (*élément constant*) are unaffected in amount.

Sources of Depot Fat.—The neutral fat in adipose tissue is derived from two main sources:

(i) From *food-fat.*

(ii) From *carbohydrate.* The classical proof (Gilbert and Lawes) is as follows: young pigs fed on barley, deposit large amounts of fat, more than could be derived from the fat and protein of the food even if it were assumed that all the carbon of the ingested protein went to form fat; clearly the fat which had been deposited must have been partly derived from carbohydrate.[2] The details of the carbohydrate→fat transformation are considered on p. 874.

It has been claimed on the basis of animal experiments that the reverse reaction can occur, *i.e.* that fat can be transformed into carbohydrate (blood glucose) on a *large scale*; these experiments are, however, open to other interpretations, and it is now generally considered that fatty acids cannot be a *major* source of blood glucose (p. 874).

Relation of Liver to Fat Metabolism.—(1) When fats are to be used in the body they are withdrawn (by unknown means) from the fat reserves, *i.e.* the adipose tissue cells, and pass to the liver; the fat content of the liver may be little altered, as the fat is broken down as fast as it arrives.

(2) The neutral fat content of the liver is *increased* in the following conditions:

(i) On a high fat diet which is also deficient in the so-called *lipotropic actors* (*i.e.* choline and methionine; see p. 867).

(ii) In starvation.

(iii) After pancreatectomy, in animals kept alive by adequate doses of insulin.[3]

The fat-laden liver arising from these conditions is called a *fatty liver.*

[1] Adipose tissue deserves more respect for its physiological activities than it customarily receives, especially as its storage function is so frowned upon by fashion.

[2] Similarly, potatoes are rightly condemned as "fattening" by those anxious about their figures, though the fat content of potatoes is almost nil (carbohydrate 19%, protein 2%, water 78%, and fat 0·1%).

[3] Fatty liver also occurs after injection of certain anterior pituitary extracts.

(3) The neutral fat in the liver is broken down by hydrolysis into *glycerol* and *fatty acids* (enzyme: *liver lipase*).

(i) The *glycerol* is utilized via the pathways of carbohydrate metabolism (p. 848).

(ii) The *fatty acids* are oxidized to fragments containing 2 C each; these fragments are either (*a*) completely *oxidized* to CO_2 and H_2O with energy liberation, or (*b*) *recombined* to give *acetoacetic acid* (a 4C compound, $CH_3.CO.CH_2.COOH$); this latter process is termed *ketogenesis*, since acetoacetic acid is a ketone.

It appears that the *liver* cannot further metabolize acetoacetic acid; any acetoacetic acid that the liver manufactures from fatty acids must be distributed to the *tissues*, where it is completely oxidized to yield CO_2, H_2O, and energy.

These reactions are more fully discussed on pp. 869–874.

(4) In circumstances of carbohydrate deficiency, the metabolism of liver fat can play a part in the maintenance of the *blood glucose* level. The liver can (i) *increase* the metabolism of fat (including the production of acetoacetic acid for energy utilization in tissues) and thus "spare" the available carbohydrate, and (ii) possibly convert some *small* part of the fat into glucose (or glycogen).

A detailed discussion is given on pp. 874, 858.

Fatty Liver. Lipotropins.—In the condition of fatty liver, a grossly enlarged liver containing massive depositions of neutral fat is found. The causes of fatty liver were enumerated in (2) above; each of the three conditions there mentioned will now be considered in detail.

(1) HIGH FAT DIET.—In animals given a high neutral fat diet, fatty liver develops. Under such circumstances, fat becomes the principal source of energy for the body and it is appropriate that large amounts of fat should be brought to the liver for complete or partial dissimilation; but one would expect that a suitable balance would be struck between the uptake of fat by the liver and its complete oxidation or its redistribution to the tissues as acetoacetic acid. It is surprising to find, therefore, that the uptake of fat by the liver should exceed, so markedly, its rate of despatch. The fat content of the fatty liver is, however, considerably *decreased* by administering: (i) *methionine* (or large amounts of proteins rich in methionine); or (ii) *choline*, or the related substance *betaine*; or (iii) *lecithin* (which contains choline).

Rôle of Lipotropins.—A substance which *reduces* the amount of liver fat is called a *lipotropin* (*lipotropic factor*). It is thought that the lipotropins other than choline itself are effective because they contain choline or because they promote choline synthesis. Considering the substances already mentioned we find:

(i) Methionine is a *methyl donor* (p. 882); it supplies labile methyl (CH_3–) groups to *ethanolamine* (cholamine, β.amino-ethanol) to form choline. The ethanolamine is synthesized in the body from glycine (probably via serine).

$$\underset{\text{glycine.}}{H_2N.CH_2.COOH} \longrightarrow \underset{\text{ethanolamine.}}{H_2N.CH_2.CH_2OH} \xrightarrow[\substack{\text{from}\\ \text{methionine.}\\ \text{or betaine.}}]{+3\,CH_3-} \underset{\text{choline (base).}}{(CH_3)_3\overset{+}{N}CH_2CH_2OH \atop OH'}$$

$$\underset{\text{methionine.}}{CH_3{-}S.CH_2.CH_2.CH(NH_2)COOH} \qquad \underset{\text{betaine.}}{(CH_3)_3\overset{+}{N}.CH_2.COO'}$$

(ii) Betaine (a methylated glycine derivative) is also a *methyl donor*, and its methyl groups can be used to synthesize choline from ethanolamine. It is unlikely that betaine can be transformed directly into choline, though choline is known to be oxidized to betaine.

(iii) Lecithin *contains* choline as part of the molecule (p. 861) and liberates choline on hydrolysis.

The mode of action of choline as a lipotropin is unknown. It may promote the conversion of liver fat into choline-containing phospholipids (*e.g.* lecithins) which are more readily transferred from the liver into the blood. The important point, however, is that *choline prevents and cures fatty liver.*

Choline deficiency also gives rise to hæmorrhagic necrosis of the *kidneys*; this disorder may likewise be due to insufficient lecithin formation in the renal cells.

Transport and Rôle of Choline.—Choline is not transported in the blood in the free state but always in the *combined* form, *i.e.* as part of phospholipide molecules (lecithin, sphingomyelin). The normal level of blood choline in man (present as phospholipide) is 30–50 mg–% ; in states of choline deficiency giving rise to fatty liver, the blood choline is much less.

Choline is an indispensable constituent of the body; if the necessary groupings are not available for its synthesis, then sufficient choline must be provided, as such, in the food. It is not strictly speaking a vitamin (though occasionally put in this category) because it can be synthesized in adequate amount by the body. As explained on p. 883, choline promotes creatine synthesis because of its action as a *methyl donor* via methionine.

(2) Starvation.—In starvation, the stress of metabolism also falls on fat, *i.e.* the fat in the depots. As lipotropins are not available in adequate amounts, fat accumulates in the liver.

(3) After Pancreatectomy.[1]—If a pancreatectomized animal is kept alive with insulin, fatty liver develops even on a fat-free diet (*e.g.* a diet of lean meat and sugar). In pancreatectomized dogs, the liver weight may increase fourfold and the total lipide content increase thirtyfold over the normal values. The condition is cured by administering large amounts of choline or of methionine. The question arises why the high meat intake (which is rich in combined methionine, *i.e.* present in the protein molecules) does not prevent the development of fatty liver. It seems that, owing to the *absence of pancreatic juice*, a factor (probably a proteolytic enzyme) is lacking which is necessary for the splitting off of the methionine from the food protein molecules in the gut; consequently, protein digestion and methionine liberation are impaired. The fatty liver can be cured (i) by administering raw pancreas or fresh pancreatic juice, both of which contain the missing factor, or (ii) by administering protein hydrolysates which contain free methionine. [In pancreatic disease with deficiency of the external secretions it is therefore advisable to give lipotropins (choline, methionine) to prevent the development of fatty liver (p. 795).] The *subcutaneous injection* of a pancreatic extract called *lipocaic* is said to cure the fatty liver which develops after pancreatectomy; lipocaic must thus act outside the bowel but its mode of action is unknown. The liability of the chronic diabetic patient to develop fatty liver should be noted (p. 924).

[1] Chaikoff and Entenman, *Advances in Enzymology*, 1948, *8*, 171

EFFECTS OF FATTY LIVER.—When the liver cells are laden with fat their functional activity is directly depressed; in addition the swollen cells compress the vascular capillaries, and so diminish the blood supply, especially to the centralobular cells (p. 821). With the development of fatty liver *hepatic function is impaired*, the most obvious symptom being *reduced glucogenesis* with the result that the fasting blood glucose level of an affected diabetic patient may fall to normal (or lower) even *in the absence of insulin*. The patient is also extremely sensitive to injected insulin (p. 927).

End results of Fatty Liver: Diffuse Hepatic Fibrosis (Cirrhosis).—Gross chronic fatty liver leads to the death of many liver cells and their replacement by scar tissue. The condition is then called *diffuse hepatic fibrosis* (*cirrhosis* of the liver). The scarring, by interfering with the hepatic blood supply, aggravates the condition, setting up a vicious circle leading to more necrosis and more scarring. Diffuse hepatic fibrosis secondary to fatty liver occurs clinically in the following states:

(i) In severe *malnutrition* (especially in the tropics). The diet is deficient in fat and in protein; the lack of lipotropins is probably the main factor responsible for the condition.[1]

(ii) In so-called "*alcoholic cirrhosis*": the liver is initially "fatty" and later shows fibrosis. The essential cause is a diet deficient in lipotropins aggravated by their defective absorption from the intestine. The drinking of excessive amounts of alcohol contributes only indirectly to the condition by leading to loss of appetite and decreased food intake and by setting up gastro-enteritis, thus interfering with absorption.

Relation of Liver to Ketogenesis.—(1) KETONE BODIES.—The name "ketone bodies" (or "acetone bodies") is applied to the following three substances, which form a metabolically-related group:

acetoacetic acid ($CH_3.CO.CH_2.COOH$)
β.hydroxy-butyric acid ($CH_3.CHOH.CH_2.COOH$)
acetone ($CH_3.CO.CH_3$)

(i) Acetoacetic acid is the parent substance of the group, and is formed during the dissimilation of (*a*) long-chain fatty acids from fat (p. 871), and (*b*) certain of the essential amino-acids (p. 888). In the body, acetoacetic acid is found associated with β.hydroxy-butyric acid and small amounts of acetone.

(ii) β.Hydroxy-butyric acid is the reduction product of acetoacetic acid (*i.e.* it is formed by the addition of 2H); the two acids are freely inter-convertible.

(iii) Acetone only arises from acetoacetic acid by spontaneous and *non-reversible* decarboxylation (loss of CO_2), a reaction which occurs chiefly in the lungs and bladder.

It should be noted that one of the three "ketone bodies," namely β.hydroxy-butyric acid, is *not a ketone*, though it is metabolically derived from one (acetoacetic acid).

The inter-relationships of the ketone bodies are summarized in Fig. 560.

The liver is the only organ which produces ketone bodies on any significant scale. A comparison of the ketone body content of the blood which enters

[1] There is an interesting tropical (mainly African) deficiency disease called *kwashiorkor* which is characterized by an enlarged grossly fatty liver and other disturbances (steatorrhœa, macrocytic anæmia, and œdema).

and leaves the liver shows that the *liver forms ketone bodies and delivers them into the hepatic vein for circulation to, and dissimilation by, the tissues.* All the tissues, with the *exception* of the *brain* and the *liver* itself, can dissimilate acetoacetic acid to CO_2, H_2O and usable energy (p. 872).

(2) BLOOD KETONE LEVEL. KETOSIS.—The normal non-fasting blood ketone level is small (0·5–2 mg–%); even a short-term fast (of two or three days) increases this level as much as fiftyfold. The amount of circulating ketone depends upon the balance between (i) *ketone formation by the liver* and (ii) *ketone dissimilation by the tissues*. Little is known about the factors that determine ketone dissimilation; it is not apparently influenced by the hormones. There is, however, a maximum amount of fat which the tissues can use (mostly as acetoacetic acid), namely about 2·5 g. of fat per kg. per day, equivalent to 175 g. of fat daily in a 70 kg. man. The rate of ketogenesis in the liver varies greatly according to circumstances (p. 875). If

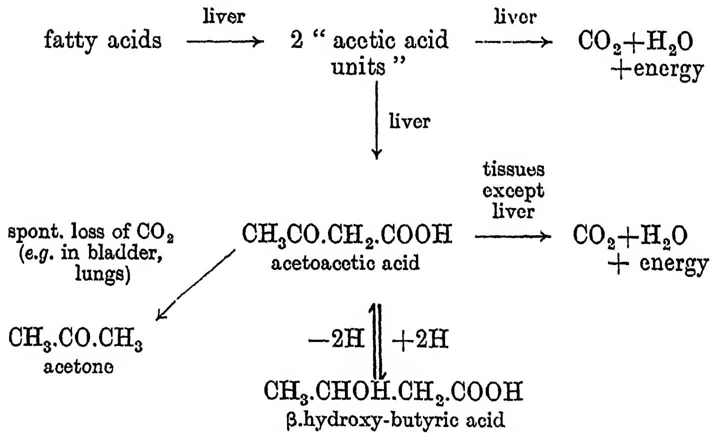


FIG. 560.—Inter-relationships of the " Ketone Bodies " [cf. with Fig. 562].

ketogenesis proceeds at an unduly high rate, exceeding the rate at which dissimilation can be carried on by the tissues, then ketones accumulate in the blood. This condition is called *ketosis*, and may lead to the excretion of ketone bodies in the urine. In extreme ketosis, the urinary ketone output may reach 100–120 g. per day.

The responses of the body to the accumulation of ketones in the blood are discussed on pp. 924, 102.

(3) MECHANISM OF KETONE BODY FORMATION. FATTY ACID OXIDATION.—The neutral fat in the liver is hydrolyzed as required, releasing glycerol and *long-chain fatty acids*, the majority containing 16 or 18 C atoms (*e.g.* palmitic, stearic, and oleic acids).

The next stage is the breakdown, mainly in the liver, of these long carbon chains into " fragments " containing two carbon atoms each; this is considered to occur in two successive steps (Fig. 561):

(i) Firstly, an *oxidation* converts every alternate $-CH_2-$ group of the chain into $-CO-$, starting from the $-CH_2-$ in the β. position relative to the

acid –COOH group; such oxidation is called *alternate β.oxidation.* In Fig. 561, the β. position is marked $\overset{*}{C}$, and oxidation occurs there and at each alternate position, marked $\overset{+}{C}$ along the chain.

(ii) The oxidation is followed by a *cleavage* of the bonds between the remaining $-CH_2-$ groups and the –CO– groups formed in step (i), giving "2C fragments." The positions of cleavage are indicated in Fig. 561 by the thin vertical lines. The addition of H_2O or H_3PO_4 is involved in this cleavage.

$$CH_3.\overset{+}{C}H_2.CH_2.\overset{+}{C}H_2.CH_2.\overset{+}{C}H_2.CH_2.\overset{+}{C}H_2.CH_2.\overset{+}{C}H_2.CH_2.\overset{+}{C}H_2.CH_2.\overset{*}{C}H_2.CH_2.COOH$$

palmitic acid

$$\Big\downarrow \text{(i) alternate } \beta \text{. oxidation}$$

$$CH_3.CO \mid CH_2.CO \mid CH_2.CO \mid CH_2.CO \mid CH_2CO \mid CH_2.CO \mid CH_2.CO \mid CH_2.COOH$$

$$\Big\downarrow \text{(ii) cleavage}$$

$$\underset{\text{unit}}{2C} + \underset{\text{unit}}{2C} + \underset{\text{unit}}{2C} + \underset{\text{unit}}{2C} + \underset{\text{unit}}{2C} + \underset{\text{unit}}{2C} + \underset{\text{unit}}{2C} + \underset{\text{unit}}{2C}$$

FIG. 561.—Breakdown of Palmitic Acid to "Acetic Acid Units."

The 16C chain of palmitic acid yields 8 "fragments," each containing 2 C atoms and derived from $-CH_2-CO-$. These fragments are thus related to (and largely behave as) acetic acid, $CH_3-CO-OH$, but are probably not identical with it; the fragments will be referred to here as "2C fragments" or as "acetic acid units" (in inverted commas).[1]

Isotope experiments using acids containing "labelled" C atoms in known positions have given the following additional information:

(i) The "2C fragments" formed in the liver can be *completely dissimilated* in the liver to CO_2, H_2O, and energy, via the citric acid cycle (p. 872).

(ii) Instead of being completely oxidized, the "2C fragments" can react together in random fashion to form molecules of the 4C ketone body, *acetoacetic acid*:

$$2 \text{ "2C fragments" (acetic acid units).} \longrightarrow \underbrace{CH_3.CO}_{\text{2C fragment.}} - \underbrace{CH_2.COOH}_{\text{2C fragment.}}$$

Thus a molecule of palmitic acid (16 C) is split in the liver into 8 highly reactive "2C fragments"; these immediately give rise to 16 molecules of CO_2, or 4 molecules of acetoacetic acid ($4 \times 4C$), or a proportion of each. The acetoacetic acid is *distributed to the tissues*, and is there oxidized to CO_2. It is usually assumed that the first step in the oxidative dissimilation of acetoacetic acid by the tissues, is its re-splitting to "2C fragments" (p. 872).

[1] As might be expected, the terminal "2C fragments" from each end of the fatty acid chain differ from each other and from the "2C fragments" from the main part of the chain; the difference has no major metabolic significance.

The proportion of "2C fragments" forming acetoacetic acid rather than being immediately oxidised to CO_2 is determined by the rate of the simultaneous carbohydrate dissimilation in the liver. If the rate of *carbohydrate dissimilation is depressed* (as in diabetes or starvation), then the proportion of *acetoacetic acid formed rises*, and ketosis may develop. A possible explanation for this effect is given on p. 875, but it must be admitted that we still know very little about the precise factors which regulate hepatic ketogenesis.

Utilization and Dissimilation of Fat.—The key step in the metabolic utilization of fat is the oxidation and cleavage of the long-chain fatty acids to yield reactive "2C fragments" ("acetic acid units") as described above. This occurs mainly in the liver (though there is evidence for some normal oxidation of long-chain fatty acids in other organs, *e.g.* a liverless dog can oxidize up to 40% of administered fatty acid).

"Acetic acid units" are highly reactive, and immediately they are formed they undergo other changes, as follows:

(1) Self-Condensation to Give Acetoacetic Acid.—The random self-condensation of two "acetic acid units" forms the 4C compound, acetoacetic acid (p. 871). This is carried in the circulation to other tissues which can dissimilate it (see 2 (i) below).

(2) Dissimilation to CO_2 and H_2O.—"Acetic acid units" and acetoacetic acid are probably completely dissimilated by entering the citric acid (Krebs) cycle, *i.e.* the set of enzyme systems which constitute the "common metabolic pathway" for most of the foodstuffs (pp. 850, 853).[1]

(i) An "acetic acid unit" ("2C fragment" from fatty acid) is metabolized in exactly the same way as the acetic acid which is formed by the oxidative decarboxylation of pyruvic acid (from carbohydrate (p. 850)). It combines with oxaloacetic acid (4C) to give the 6C tricarboxylic acid, citric acid; this then "goes round the cycle," losing two carbon atoms in the form of two molecules of CO_2 by two successive decarboxylations. As a result, energy is liberated which is made available to the body through the high-energy phosphate bonds of ATP (p. 852); oxaloacetic acid is regenerated, and is ready for reaction with a further "acetic acid unit" (Figs. 562, 558).

(ii) It appears that the main step in the dissimilation of *acetoacetic acid* in the tissues is its re-splitting into two "2C fragments"; these are then oxidized by the enzymes of the Krebs citric acid cycle present in the tissue concerned. Liver and brain seem unable to split acetoacetic acid, and thus cannot utilize it to any extent.

Acetoacetic acid may also be able to enter the citric acid cycle directly (say at a 4C stage) but this must constitute a minor pathway.

It can be seen from Fig. 562 that for the complete dissimilation of fats via "2C fragments," oxaloacetic acid must be available from some source (*e.g.* from pyruvic acid (p. 850) or from aspartic acid (p. 889)); otherwise the "2C fragments" will form correspondingly larger quantities of acetoacetic acid, and *ketosis* will develop (p. 875).

[1] "Acetic acid units" (2C) and acetoacetic acid (4C) are the "ready money" of fat metabolism. The long C chain of the fatty acids is appropriate for storage purposes only, and represents capital locked away in securities. This wealth is not negotiable directly, but must be converted into the ready money of 2C and 4C fragments (2C fragments for buying a paper or paying a bus fare; 4C "notes" for putting the wealth into circulation). Just as ready money in sufficient quantity can be converted into securities, so can 2C fragments from any source be built up into the long C chain of the fatty acids (p. 873).

(3) ACETYLATION REACTIONS.—The active "acetic acid unit" can be employed by the body in acetylating reactions, *e.g.* the synthesis of acetylcholine (p. 508) or the acetylation of amines during their "detoxication" by the liver (p. 828).

$CH_3.CO{-}O.CH_2CH_2\overset{+}{N}(CH_3)_3OH'$ acetylcholine

$CH_3.CO{-}NH.C_6H_4.SO_2NH_2$ acetylsulphanilamide

(4) USE AS BUILDING UNIT FOR BODY COMPONENTS.—Evidence from isotope experiments shows that labelled C atoms from acetic acid (or compounds giving rise to "acetic acid units") can be built into the molecules of (i) *cholesterol* (most of its 27 C atoms come directly from "acetic acid units"); (ii) *hœm porphyrins*, and (iii) the intermediates of the citric acid cycle and their derivatives.

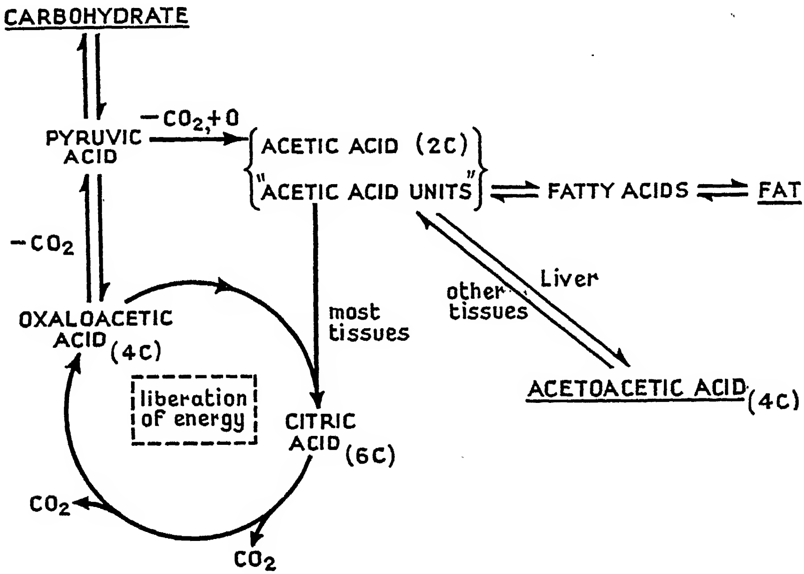


FIG. 562.—The Formation and Dissimilation of Fat. (Diagram by J. B. Jepson.) [Note the irreversibility of the pyruvic acid→acetic acid step.]

(5) RESYNTHESIS OF FATTY ACID.—"Acetic acid units" (from any source) can be built up into long-chain fatty acids, and thence into fat. This presumably occurs by the successive condensation of "2C fragments" followed by, or simultaneous with, a reduction process (*i.e.* the reversal of the breakdown of fatty acids), and requires an external supply of energy. In support of this view we have:

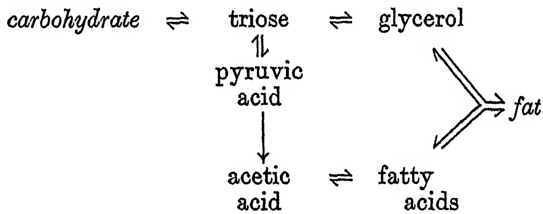
(i) all fats are built from fatty acids with an *even* number of C atoms;

(ii) isotope experiments show that when acetic acid $CD_3.^{13}COOH$, labelled with ^{13}C and deuterium (D, heavy hydrogen), is fed to animals, the liver fats have an isotope distribution in the fatty acid chains thus: $-CD_2.^{13}CH_2{-}CD_2.^{13}CH_2{-}CD_2.^{13}COOH$.

Coenzyme-A.—The presence of a *coenzyme* is required for reactions involving "acetic acid units"; this coenzyme is termed *coenzyme-A* (A for Acetylation) and contains *pantothenic acid*—yet another example of a *B*-group vitamin with a specific rôle in an enzyme system (p. 1024 ff.).

Integration of Fat and Carbohydrate Metabolism.—CONVERSION OF CARBOHYDRATE TO FAT.—The well-known conversion of dietary carbohydrate into depot fat has been mentioned on p. 866. It is obvious from Fig. 562 that the pathways of carbohydrate and fat metabolism meet at the common intermediate, acetic acid (or "2C fragment"). Since fatty acid is reversibly formed from "2C fragments" the way is open for the transformation of carbohydrate via pyruvic acid and acetic acid into *fatty acids*. This transformation is promoted by insulin and depressed by anterior pituitary hormones (pp. 913, 920).

Glycerol from the hydrolysis of fats is converted into a triose (probably phosphoglyceraldehyde) and joins the main pathway of carbohydrate metabolism (p. 848). This reaction is reversible, allowing carbohydrate to be used for the synthesis of *glycerol* when this is needed for the deposition of fatty acids as body fat.



CONVERSION OF FAT TO CARBOHYDRATE.—As stated above, the *glycerol* of fats can join the reversible pathway of carbohydrate metabolism and thus be built into glucose or glycogen. This would allow the formation of only 12 g. of blood glucose from 100 g. of fat, whereas it has been thought that much greater conversion than this could occur in the fasting animal. There is no doubt (from isotope experiments) that C *atoms* from acetic acid (and other molecules giving rise to "acetic acid units," *e.g.* fatty acids) can be built into the molecule of glucose by the liver. A consideration of Figs. 562, 559, however, shows that *no net increase* of glucose could arise in this way. There is no known pathway in animals by which the fatty acid→glucose conversion could be accomplished. Any net increase in the glucose of the body fluids that *does come directly* from fatty acid must be derived from the synthesis of, say, succinic acid (4C) which would *directly* enter into the Krebs cycle and could thus be converted into oxaloacetic acid and thence into pyruvic acid and glucose; this must be a minor pathway. Under conditions of carbohydrate deficiency, the chief function of the liver is to *provide acetoacetic acid for utilisation by the tissues, thus "sparing" the utilisation of the more valuable blood glucose.*

DEPENDENCE OF FAT METABOLISM ON CARBOHYDRATE.—The complete dissimilation of fat via "acetic acid units" requires available supplies of oxaloacetic acid to "catalyse" the Krebs cycle (Fig. 562). This must be supplied from a source *other* than fat (*since the pyruvic acid→acetic acid reaction is irreversible*); the chief source of oxaloacetic acid is carbohydrate,

via pyruvic acid and the carboxylase reaction (p. 850). *The complete dissimilation of fats in the liver thus requires the simultaneous oxidation of carbohydrate.*[1] Any change which reduces the rate of carbohydrate oxidation in the *liver* will reduce the rate of complete oxidation of " acetic acid units " locally, without reducing their rate of formation; indeed, their rate of formation may *increase* if fat is the only available energy source. However, as these highly reactive " 2C fragments " cannot accumulate, they undergo self-condensation to form acetoacetic acid. The lower the rate of carbohydrate utilization in the liver relative to the fat utilization, the greater is the production of acetoacetic acid, and the amount which circulates in the blood correspondingly increases. If the tissues cannot dissimilate the increased supplies of acetoacetic acid which reach them, then *ketosis* results. Starvation, diabetes, and the other circumstances that lead to ketosis all involve a decrease in the normal ratio of carbohydrate to fat utilization in the liver.

Causes of Ketosis.—Ketosis occurs clinically and experimentally in the following conditions:

(1) With a *low glycogen and a high fat content in the liver, e.g.* on a high fat, low carbohydrate diet, or in starvation (when the liver glycogen reserves are exhausted and depot fat is being utilized). It is thought in a vague way that under these conditions fat becomes the " principal substrate of the liver cell " leading to excessive ketone body formation (p. 870). It should be noted that a small proportion of the ketones may be derived from ketogenic amino-acids (p. 888). The formation of glucose from glucogens may be simultaneously depressed. As the tissues cannot cope with all the ketones supplied to them, the blood ketone level rises.[2]

(2) From the injection of certain *anterior pituitary extracts* in normal animals (p. 937). These extracts " mobilize " depot fat, which passes into the blood stream to reach the liver; liver fat is thus increased. Glucose dissimilation is simultaneously depressed, and liver glycogen is decreased owing to its conversion into blood glucose. For this reason, or because of some additional *specific* action on the liver, these anterior pituitary extracts increase ketone formation in the liver, with resulting ketosis.

(3) *Ketosis of Diabetes Mellitus. Action of Insulin.*—After removal of the islet tissue of the pancreas, the unantagonized action of the anterior pituitary leads to ketosis as described in (2) above. Injected insulin (p. 913) acts in the opposite manner: it promotes deposition of depot fat and decreases the flow of fat to the liver; it increases glucose dissimilation and glycogen deposition in the liver; it possibly annuls the specific ketone-producing action of the anterior pituitary. Thus by its action on the fat depots and on the liver, insulin abolishes the ketosis of diabetes mellitus; it has no effect on ketone utilization by the tissues.[3]

[1] This idea used to be expressed crudely but vividly in the phrase " The fats burn in the flame of the carbohydrates."

[2] People differ greatly in their susceptibility to ketosis. Thus Eskimos can tolerate high fat diets that would cause gross ketosis in the average European. The body can apparently be " trained " gradually in the utilization of very high fat diets.

[3] It is found empirically in clinical cases of diabetes mellitus that ketosis is avoided when the fat content (F) of the diet does not exceed the sum of twice the carbohydrate (C) and half the protein (P) contents. Briefly, F should be less than $2C+\frac{1}{2}P$, all measured in grammes.

PROTEIN METABOLISM

Proteins.[1]—The proteins have complex molecular structures built mainly from *a-amino-acids* linked together in chains. The linkage between the amino-acids is called a *peptide bond*; molecules built up from many amino-acids are called *polypeptides*; proteins are types of polypeptide, but their structures are so large and complex that relatively little of their molecular architecture has yet been elucidated. Hydrolysis (digestive or otherwise) converts a protein through stages of intermediate complexity (conventionally called metaproteins, proteoses, and peptones) to a mixture of amino-acids. About twenty different amino-acids have been found in the various proteins studied; most proteins contain a selection of 15–18 of these, though some proteins are built from only a few different amino-acids.

The amino-acids have the general formula R.CH(NH_2).COOH, where R is any one of a variety of organic groupings as shown below. The peptide bonds of a polypeptide chain are —CO—NH—(amide) linkages of the following type:

$$- \; - \; —CO—NH.\overset{\overset{R'}{|}}{C}H.CO—NH.\overset{\overset{R''}{|}}{C}H.CO—NH— \; - \; -$$

peptide bonds

The physiological properties of the free amino-acids depend upon the nature of their constituent R groups; the properties of the proteins are governed by the nature and mutual arrangement of the constituent amino-acids.

Amino-Acids.—The names and formulæ of the principal amino-acids obtained by the hydrolysis of proteins are set out below:

Amino-acids with unsubstituted C chains:

1. Glycine (α-amino-acetic acid)
   $NH_2.CH_2.COOH$ (R=H; the simplest amino-acid)
2. Alanine (α-amino-propionic acid)
   $CH_3.CH(NH_2)COOH$ (3C chain)
3. Valine
   $(CH_3)_2CH.CH(NH_2)COOH$ (5C branched chain)
4. Leucine
   $(CH_3)_2CHCH_2.CH(NH_2)COOH$ (6C branched chain)
5. Isoleucine
   $(CH_3)(C_2H_5)CH.CH(NH_2)COOH$ (6C branched chain)

Hydroxyl-substituted amino-acids:

6. Serine (α-amino-β-hydroxy-propionic acid)
   $HO.\underset{\beta}{CH_2}.\underset{\alpha}{CH}(NH_2)COOH$ (3C chain)

[1] Sahyun, *Outline of the Amino-Acids and Proteins*, 2nd edn., New York, 1948. "Amino-Acids and Proteins," *Cold Spring Harbor Symp. Quant. Biol.*, 1949, XIV. Haurowitz, *Chemistry and Biology of Proteins*, New York, 1950.

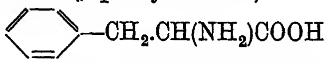
7. Threonine
   $CH_3CH(OH).CH(NH_2)COOH$ (4C chain)

Sulphur-containing amino-acids :

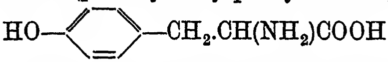
8. Cysteine (α-amino-β-thiol-propionic acid)
   $HS.CH_2.CH(NH_2)COOH$ (3C chain—note the reactive HS. group)
9. Cystine (dicysteine, the oxidation product of cysteine)
   $S.CH_2.CH(NH_2)COOH$
   |
   $S.CH_2.CH(NH_2)COOH$
10. Methionine (α-amino-γ-methylthio-butyric acid)
   $CH_3S.\underset{\gamma}{CH_2}CH_2.\underset{\alpha}{CH}(NH_2)COOH$ (4C chain—note the CH_3S. group)

Aromatic amino-acids, derived from alanine (2.) *:*

11. Phenylalanine (β-phenyl-alanine)

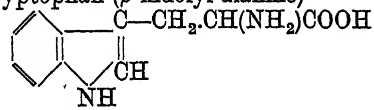


12. Tyrosine (*para*-hydroxy-phenyl-alanine)



(a *phenol*)

13. Thyroxine, an iodine-containing derivative of tyrosine, see p. 975.
14. Tryptophan (β-indolyl-alanine)



The amino-acids (1 to 14) set out above are *neutral* substances, since they contain one NH_2. (basic) group and one .COOH (acidic) group which mutually neutralise each other.

Acidic amino-acids :

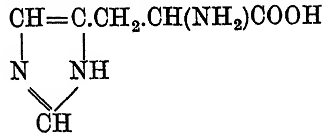
15. Aspartic acid (α-amino-succinic acid) Asparagine= the β-amide (4C chain)
   $CH_2.COOH$
   |
   $CH(NH_2)COOH$
16. Glutamic acid (α-amino-glutaric acid) Glutamine= the γ-amide (5C chain)
   $CH_2CH_2.COOH$
   |
   $CH(NH_2)COOH$

The two amino-acids (15 and 16) set out above contain an additional COOH (acidic) group in the **R** group (*i.e.* they are *mono*amino-*di*carboxylic acids) and are *acidic* substances :

Basic amino-acids :

17. Arginine (α-amino-δ-guanidino-valeric acid)
   $$NH_2—\overset{\overset{NH}{\|}}{C}—NH.\overset{\delta}{C}H_2CH_2CH_2.\overset{\alpha}{C}H(NH_2)COOH$$ (5C chain)
18. Lysine (α : ε-diamino-caproic acid)
   $NH_2.CH_2CH_2CH_2CH_2.CH(NH_2)COOH$ (6C chain)

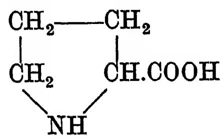
19. Histidine (β-iminazolyl-alanine)



The three amino-acids (17, 18, 19) set out above have nitrogen-containing R groups and are *basic* substances:

Imino-acid:

20. Proline



(5C chain)

Substance 20 is actually an *imino-acid*, with an α-NH as part of a ring system. It is *neutral.*

Other amino-acids of metabolic significance are known, but they have not been found in protein molecules. Examples are *ornithine* and *citrulline* of the Krebs urea cycle (p. 888, Fig. 568), and *homocysteine* formed from methionine (p. 883, Fig. 564).

Nucleoproteins. Nucleic Acids.[1]—Nucleoproteins are conjugated proteins, consisting of a basic protein combined with a *nucleic acid.* The protein moiety can be removed by acid treatment, freeing the nucleic acid.

Nucleic acids are *polynucleotides* whose exact structure has not yet been determined. On hydrolysis they yield mixed *nucleotides* with the following type of structure:

Nitrogen-containing Base—Pentose—Phosphate.

The pentose sugar is either *ribose* or *desoxyribose.* The base is either a *purine* (*adenine* or *guanine*) or a *pyrimidine* (*cytosine, thymine* or *uracil*) (formulæ: Figs. 573, 574). Hydrolysis of a nucleotide removes the phosphate leaving a *nucleoside* (*i.e.* Nitrogen-containing Base—Pentose).

Nucleoproteins are important constituents of the nucleus and cytoplasm of all cells; chromosomes are constructed largely from nucleoprotein.

Several of the vitamins of the *B*-group are built up by the body into nucleotides, and in this form function as coenzymes for fundamental dissimilation reactions (p. 1024 ff.). The nucleotides adenosine-diphosphate and -triphosphate (ADP, ATP) are of great metabolic importance as carriers of "high-energy phosphate bonds" (p. 842, Fig. 553).

Digestion of Proteins.—No digestion of protein occurs in the *mouth.*

Stomach.—(i) The pepsin-HCl mixture acts on proteins, converting them through stages of decreasing complexity (metaprotein, primary and secondary proteoses) to peptones (large polypeptides). (ii) Caseinogen is converted by

[1] "Nucleic Acids and Nucleoproteins," *Cold Spring Harbor Symp. quant. Biol.*, 1947, *12.* "Nucleic Acid," *Symp. Soc. exp. Biol.*, 1947, *1.* Davidson, *The Biochemistry of the Nucleic Acids*, London, 1950.

rennin into soluble casein. In the presence of calcium salts, casein is precipitated as insoluble calcium caseinate, and is subsequently digested to the peptone stage.

Duodenum.—Trypsin and chymotrypsin of pancreatic juice, acting in an alkaline medium, rapidly break down protein through intermediate stages to peptones and small peptides and may liberate some amino-acids. Their digestive action is facilitated by the bile-salts. It must be remembered that in the intestine these enzymes are acting on protein which has already been partially digested by pepsin.

Small Intestine.—Succus entericus contains the enzyme mixture " erepsin," a system of peptidases which specifically hydrolyse terminal peptide bonds, and thus completes the breakdown of polypeptides to amino-acids ; this occurs partly in the lumen and partly in the wall of the intestine.

Digestion of Nucleoprotein.—The protein portion is removed from nucleoprotein by the acid of the stomach, and digested with the other food proteins. The freed nucleic acid is hydrolysed enzymatically in the gut to nucleotides, nucleosides, and finally to the constituent pentoses, purines, and pyrimidines. These are then absorbed and metabolised. The metabolism of purines, pyrimidines and nucleic acids is discussed on p. 895.

Absorption of Amino-Acids.—The amino-acids, which are the end-products of protein digestion are absorbed into the intestinal capillaries and thence enter the portal vein ; absorption is in part due to simple diffusion owing to the higher amino-acid concentration in the intestinal lumen, but active processes may also be involved. There is evidence that, in *special circumstances,* certain proteins and peptides may be absorbed from the gut without previous *complete* hydrolysis to amino-acids (*e.g.* during the digestion by an animal of its own plasma protein).

Amino-Acid Pool. Amino-Acid Turnover. Protein Synthesis.—The amino-acids after absorption into the blood diffuse throughout the body fluids and so reach all the tissue cells. At the same time most of the tissue proteins (both " structural " protein and " functional " enzyme protein) are continually undergoing disintegration to release amino-acids which likewise enter the circulation and thus become part of what is called the general " amino-acid pool." This steady and rapid tissue protein breakdown is taking place on a large scale (see below). No functional distinction can be drawn between the fate of the amino-acids derived from the food and those derived from the tissues. From the " common amino-acid pool," amino-acids are taken up by the cells (each cell according to its specific needs) to be built into the cell structure as required. If a cell takes up as much amino-acid as it loses, it is in a state of " dynamic equilibrium " ; if the loss is greater the cell wastes ; if the gain is greater the cell grows.

From the rate of incorporation of isotopically labelled amino-acids, the rate of synthesis of proteins can be calculated. In experimental animals this protein turnover is greatest in intestinal mucosa, followed by kidney, liver, brain, and muscle in that order. In *man,* the protein " turnover " involves the breakdown and re-synthesis of 80–100 g. of tissue protein per day, about half of it occurring in the liver ; on an average the plasma proteins are completely replaced every 15 days.[1] Nothing is known of the mechanism of protein synthesis in the body.

[1] Borsook, *Physiol. Rev.*, 1950, *30*, 206.

The amino-acid pool has no anatomical reality, but represents the *available* amino-acid building units. Thus the pool "contains" not only all the free amino-acids of the blood, extracellular and intracellular fluids, but also the amino-acids which would be freed by the *net* breakdown of a portion of the tissue proteins. The pool is constantly undergoing *depletion* because (i) large-scale deamination of presumably surplus amino-acids takes place, the amino (NH_2) groups which are split off being transformed mainly into urea, leaving a "non-nitrogenous residue"; (ii) amino-acids and their derivatives (*e.g.* creatinine) are lost in the urine and other excretions; (iii) amino-acids are continually being built up into those proteins (*e.g.* hair) which are not part of the dynamic systems. On the other hand, the amino-acid pool is always being *re-established* by amino-acids derived from the following: (i) reamination of certain non-nitrogenous residues; (ii) amination of appropriate "fragments" which are present in the common metabolic pool (and therefore derived from carbohydrate and fat breakdown); (iii) amino-acids split off from dietary protein and absorbed from the intestine into the blood. This state of affairs is termed the "continuing metabolism" of the amino-acids.

The *non-nitrogenous residues* which are left after deamination of the amino-acids are used in ways to be described (p. 887).

It should be remembered that it is not only the proteins of the body that are in a state of "dynamic equilibrium," *i.e.* a balance between simultaneous breakdown and synthesis; the same holds for practically all the materials of the body, even for such seemingly chemically inert material as the depot fat.[1]

Example of Amino-Acid Utilisation.—The general principles of amino-acid utilisation which are discussed in detail later, are well illustrated by a study of the fate of ingested *glycine*, in which both C atoms and the N atom were "labelled" with isotopes.[2] After absorption, the labelled glycine mixes with the glycine already present in the body fluids. The changes undergone by the glycine have proved to be as follows:

(i) Some is incorporated into the tissue and plasma proteins, at rates varying with each protein.

(ii) Some is built into non-protein compounds which contain glycine as part of their structure, *e.g.* glutathione, glycocholic acid (p. 798), hippuric acid (p. 828).

(iii) Some is converted reversibly into the amino-acid *serine*, and thence into *cysteine* (p. 882).

(iv) Some is used in the synthesis of other body constituents for which glycine is a *specific* precursor, *e.g.* creatine (p. 893, Fig. 572), haem and bile pigments (pp. 173, 188), purines (p. 897), choline (probably via serine, p. 867).

(v) The remainder is *deaminated*. The NH_3 split off is (*a*) converted into and excreted as *urea*, or (*b*) used to aminate various keto-acids to form other amino-acids (p. 885), or (*c*) indirectly excreted as ammonium ions (NH_4^+) in the urine.

(vi) The non-nitrogenous residue which results from the deamination

[1] "Life is a dynamic equilibrium in a polyphasic system" (F. G. Hopkins, 1913). For the evidence see Schönheimer, *The Dynamic State of Body Constituents*, London, 1942.

[2] Rittenberg, *Cold Spring Harbor Symp. Quant. Biol.*, 1948, *13*, 173. Bentley, *Ann. Reps. Chem. Soc.*, 1948, 239.

reaction (v) is either (*a*) *dissimilated* to CO_2 and H_2O, yielding energy, or (*b*) built up into *glycogen* or *glucose* (p. 888).

Essential Amino-Acids.—The term *essential* is applied to those amino-acids needed for replacement and growth, but which *cannot be synthesised* by the body in amounts sufficient to fulfil its normal requirements. They must therefore be supplied *in the diet*, usually combined in proteins. A more appropriate name for them would be the *indispensable* amino-acids, as they only differ from other amino-acids in being indispensable components of the *diet*.

The rest of the amino-acids (the *non-essential*, or *dispensable*, amino-acids) can be synthesised in the body either (*a*) by the amination of appropriate non-nitrogenous fragments derived from other sources, or (*b*) in special cases directly from the essential amino-acids. But it should be emphasised that the body is "spared the trouble" of this synthesis if the dispensable amino-acids are also available from the diet.

The essential amino-acids in adult *man* are: *valine*, *leucine*, *isoleucine*, *threonine*, *methionine* (which can also be converted to cysteine), *phenylalanine* (which can also be converted to tyrosine), *tryptophan*, *lysine*. *Arginine* and *histidine* are probably indispensable in children for growth.

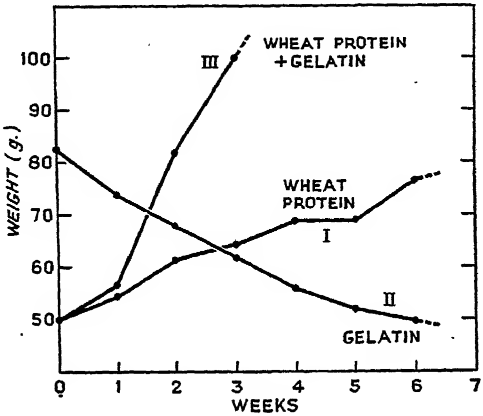


FIG. 563.—Influence of Lysine Deficiency on Growth. Growth Curves for Rats on Non-Nitrogenous Basal Diets supplemented with the various Proteins indicated. (Drawn by J. B. Jepson from data in papers by Hawk, and Osborne and Mendel).

Curve I: Wheat protein (deficient in lysine).
Curve II: Gelatin (deficient in many amino-acids, though high in lysine).
Curve III: Wheat protein + gelatin (sufficient of all the essential amino-acids). Normal growth curve.

Animals given a basal diet which contains no protein or amino-acid, but which is otherwise complete in all respects, rapidly die; if, however, all the above named amino-acids are added in appropriate amounts (either as protein or as pure synthetic amino-acids), normal health, protein replacement, and reproductive power are maintained in adult animals, and satisfactory growth occurs in young animals (Fig. 563). Rose is conducting long-term experiments with healthy men fed on diets devoid of nitrogeneous compounds except for known added amounts of pure synthetic amino-acids.[1] By measuring the nitrogen-balance (p. 901) on various amino-acid supplements, he has found that the above eight amino-acids are indispensable for human adults under normal conditions; exclusion of *any one* of these essential amino-acids leads to a negative nitrogen-balance (p. 903, Fig. 578), fatigue, loss of appetite, and "nervous irritability"; when the missing amino-acid is added to the diet perfect health is promptly restored.

[1] Rose, *Fed. Proc.*, 1949, *8*, 546; *J. biol. Chem.*, 1950, *182*, 541.

Under normal conditions, the amount of each essential amino-acid which must be supplied per day, when *all other* amino-acids are present in abundance, is of the order 0·5–2 g. The requirement for essential amino-acids varies with the amount of "strain" placed upon the synthetic capacity of the body. Thus during growth or lactation a higher intake of essential amino-acids (as mixed protein) is required (p. 1059); in addition, as mentioned above, the amino-acids arginine and histidine may become indispensable under these conditions, though they are dispensable in the normal adult, who can make them on a small scale.

Specific Metabolic Rôles of Amino-Acids.—In addition to being the building units of all the tissue proteins (including the enzymes and many of the hormones), amino-acids have special rôles in the formation of other components of the body. Some illustrative examples will be given :—

(1) Glycine is a fundamental building unit from the common metabolic pool, and its reactions are outlined on p. 880.

(2) Arginine is part of the cycle which is responsible for *urea* formation (p. 887), and provides an amidine group for *creatine* synthesis (p. 893, Fig. 572).

(3) Histidine is the precursor of *histamine* (p. 335).

(4) Aromatic Amino-Acids: *Phenylalanine* (which is indispensable) can be irreversibly converted to *tyrosine*; the latter is therefore dispensable if sufficient phenylalanine is supplied. Tyrosine is the precursor for *thyroxine* (p. 974), *nor-adrenaline*, and thence *adrenaline* (p. 729), and the dark *melanin pigments* of the hair and skin.

(5) Sulphur-Containing Amino-Acids.—Methionine, cysteine, and cystine are the sole sources of sulphur (apart from traces of the S-containing vitamins, aneurin and biotin) that can be used for synthetic purposes in the body, *e.g.* for the formation of organic sulphates (p. 900) or taurine (p. 798).

Methionine (which is indispensable) can be irreversibly converted into *cysteine* (which is therefore dispensable). This transformation occurs through the following stages (Fig. 564):

(i) Methionine forms homocysteine by transference of its CH_3 elsewhere (*e.g.* to form creatine, see (6) below, and Fig. 565).

(ii) Homocysteine (4C) combines with the 3C amino-acid serine to form the 7C compound *cystathionine* (with loss of H_2O).

(iii) Cystathionine undergoes cleavage at the homocysteine linkage of the S (with the addition of water) forming cysteine and (probably) homoserine.

It is clear that only the S atom of the original methionine appears in the cysteine. This has been proved by the use of isotopic tracers.

Cysteine is readily converted to *cystine* under oxidative conditions, and the reaction is reversed under reductive conditions. The biological action of a protein containing a free SH. group (*i.e.* combined cysteine) is often completely altered upon oxidation to the corresponding disulphide (*i.e.* combined cystine).

(6) Methionine (Biological Methylation).—Certain compounds of the body, with structures containing methyl groups (CH_3—) attached to an atom *other than* C can take part in enzymic reactions whereby these methyl groups are transferred to suitable "acceptors" which have no methyl group. Such reactions are termed *transmethylation* reactions, and the substrates (*methyl donors*) are said to possess *biologically labile methyl groups*. The most important compounds with biologically labile methyl groups are

methionine (containing CH_3—S—, Fig. 565, *choline* (containing $(CH_3)_3N^+$—, Fig. 565), and the oxidation derivative of choline, *betaine* (p. 867). The processes of transmethylation in the body have been investigated by the use of isotopes,[1] and the results are summarised in Fig. 565.

The following points should be noted :

(i) The reaction **methionine** $\underset{+CH_3,\ \ -H}{\overset{-CH_3,\ \ +H}{\rightleftharpoons}}$ **homocysteine** is reversible. However, homocysteine can only take up a methyl group to re-form methionine if the methyl group is donated by *choline* (*i.e.* the reverse of (ii) *infra*).

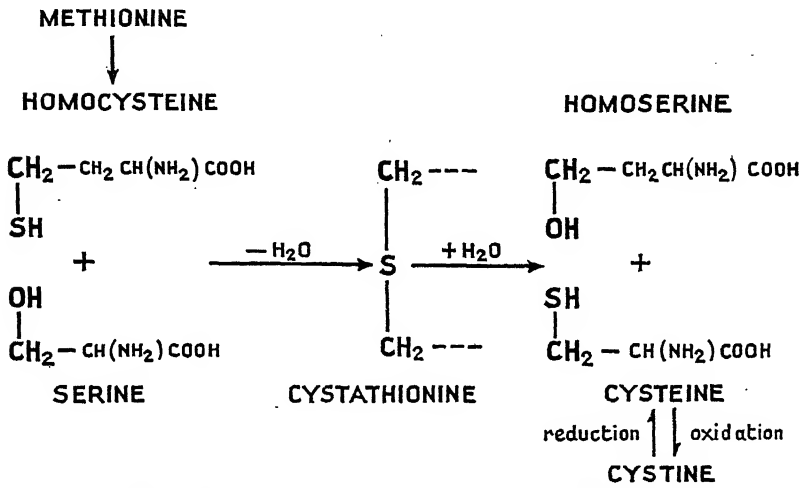


FIG. 564.—Inter-relationships of the Sulphur-containing Amino-acids.

(ii) Methionine enables choline to be synthesised in the body from ethanolamine ($NH_2.CH_2CH_2OH$) by donating the necessary three methyl groups (p. 867).

(iii) Methionine can yield methyl groups to other suitable recipients, but these reactions appear to be irreversible. Two examples are given :

(*a*) Methionine contributes a methyl group to guanidino-acetic acid, thus bringing about the synthesis of *creatine* (p. 893, Fig. 572). This methyl group must be donated directly by methionine (but may come indirectly from choline donating a methyl group to homocysteine to form methionine).

(*b*) Methionine contributes a methyl group to nor-adrenaline for the synthesis of *adrenaline* (p. 729).

These reactions of methionine illustrate how artificial and misleading is the customary division of general metabolism into subsections called carbohydrate, fat, and protein metabolism. Methionine, derived from protein, is responsible for creatine formation, and thus indirectly for carbohydrate utilisation in muscle (p. 429) ; it is responsible for choline formation, and thus indirectly with preventing undue fat accumulation in the liver (p. 867) ; it is

[1] Du Vigneaud, *Harvey Lectures*, 1942–43, 1.

(or, on a minor scale, as ammonium ions), or used in the synthesis of other amino-acids, as described below. The *non-nitrogenous residues*, after undergoing various transformations, enter the " common metabolic pool," and are either completely dissimilated or are converted into glucose, fat, ketone bodies, or other amino-acids (p. 887).

(1) *Oxidative Deamination.*—The overall reaction is the transformation of an amino-acid, $R.CH(NH_2).COOH$, to the corresponding keto-acid, $R.CO.COOH$. This involves an oxidation (or more exactly, a dehydrogenation) (enzymes: *amino-acid oxidases*; hydrogen carriers: *coenzyme-I* or *flavin* (pp. 840, 854)) to give an *imino*-acid, followed by a hydrolysis liberating ammonia, thus:

$$\underset{\text{alanine}}{CH_3.CH(NH_2)COOH} + \underset{\substack{\text{coenzyme}\\\text{(H-carrier)}}}{\text{Co-I}} \longrightarrow \underset{\alpha\text{-imino-acid}}{CH_3.C(:NH)COOH} + \underset{\substack{\text{reduced}\\\text{coenzyme}}}{\text{Co-I.2H}}$$

$$CH_3.C(:NH)COOH + H_2O \longrightarrow \underset{\text{pyruvic acid}}{CH_3.CO.COOH} + NH_3$$

(2) *Transamination.*—Deamination of an amino-acid may be coupled with the simultaneous amination of a keto-acid. This process is called *transamination* (enzymes: *transaminases*), and the result is the transference of the NH_2. group from one amino-acid and its use to synthesise another, thus:

$$\underset{\text{alanine}}{\begin{array}{l}CH_3\\|\\CH(NH_2)COOH\end{array}} + \underset{\substack{\alpha\text{-keto-}\\\text{glutaric acid}}}{\begin{array}{l}CH_2CH_2COOH\\|\\CO.COOH\end{array}} \rightleftharpoons \underset{\text{pyruvic acid}}{\begin{array}{l}CH_3\\|\\CO.COOH\end{array}} + \underset{\text{glutamic acid}}{\begin{array}{l}CH_2CH_2COOH\\|\\CH(NH_2)COOH\end{array}}$$

Transamination is *reversible* and plays an important part in both the breakdown of amino-acids and their synthesis from non-protein sources (*e.g.* from keto-acids of the citric acid cycle, p. 887).

(3) *Amination of Non-nitrogenous Residues.*—When ammonium salts or amino-acids containing amino groups " labelled " with isotopic ^{15}N are administered, the ^{15}N label is later found incorporated into the amino groups of many different amino-acids of the proteins.

Appropriate non-nitrogenous residues must therefore have been converted into amino-acids by taking up NH_2. from NH_4^+ or other amino-acids, *i.e.* by processes of direct amination or transamination. Amino-acids from the amino-acid pool are continually being broken down by deamination, and the processes of direct amination or transamination are used to resynthesise some of these amino-acids. These synthetic processes are increased if a need arises for additional quantities of amino-acid, *e.g.* the provision of glycine for " detoxication " mechanisms (p. 828). Products of deamination formed at one site can be reaminated elsewhere and so re-enter the " amino-acid pool."

(4) *Ammonia*, which is formed by the *kidney*, probably mainly from glutamine, is excreted into the lumen of the renal tubule and used in the regulation of the blood reaction (p. 97).

Urea Formation.—The surplus ammonia which is formed by deamination and not used for reamination is converted into urea. Ammonia is toxic and its toxic action is not entirely due to its strong basicity; it therefore cannot be allowed to accumulate. Urea, on the other hand, is harmless even in very high concentrations.

Urea formation occurs exclusively in the liver. Removal of the liver produces the following results (Figs. 566, 567):

(i) When urinary secretion is maintained in the liverless animal, there is a steady decrease in the blood urea, *e.g.* from 30 to 6 mg. per 100 c.c.; the amount eliminated in the urine likewise shows a progressive and marked decrease. Practically all the urea which is secreted after hepatectomy can be accounted for by the decrease in the urea content of the blood fluids.

(ii) If anuria follows the operation, the blood urea remains quite constant.

(iii) If the kidneys are removed in a normal animal there is a rise in the blood urea; if hepatectomy is now performed, the blood urea remains at the high level previously attained (Fig. 567).

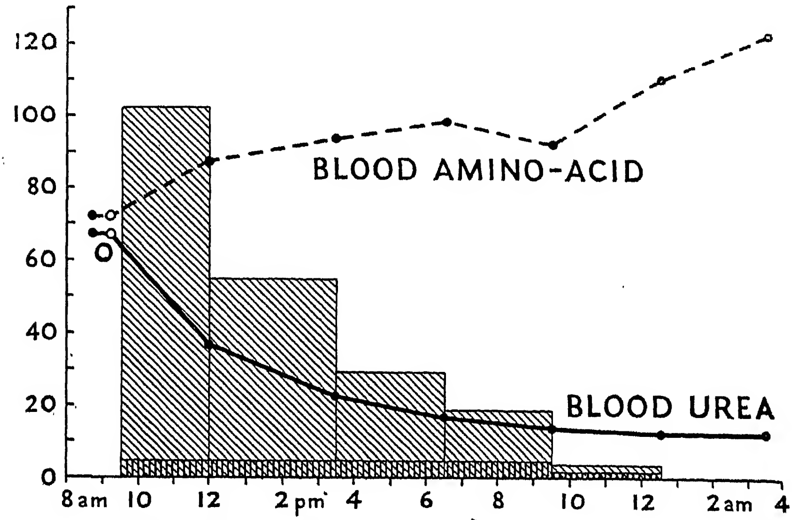


FIG. 566.—Effect of Removal of Liver on Blood Urea, Blood Amino-Acid and Urea and Amino-Acid Excretion in Urine.

Experiments on dogs. At point marked O remove liver.
Curve with broken line: blood amino-acid nitrogen in mg./*litre*.
Curve with continuous line: blood urea nitrogen in mg./*litre* (cf. Fig. 567).
Rectangles with angle hatching: urea nitrogen + ammonia nitrogen excreted in urine in mg./hour.
Rectangles with vertical hatching: amino-acid nitrogen in urine in mg./hour.
Note rise of blood amino-acid, fall of blood urea and decrease and ultimately disappearance of urea from urine. (Redrawn from Mann *et al.*, *Amer. J. Physiol.*, 1926, *78*, 259.)

These results prove that the *liver is the only site of urea formation*, and that urea once formed is not destroyed in the body.

(iv) The liver is the chief site of deamination of amino-acids. In the liverless animal the blood amino-acid content progressively rises (Fig. 566); injected *amino-acid is not deaminated* to the normal degree and does not give rise to extra urea.

Essentially, urea formation in the liver consists in the union of NH_3 (2 mols.) and CO_2 (1 mol.) with the elimination of H_2O. This, however, does not take place directly but through a cyclical system of enzyme reactions

(called the *Krebs urea cycle*; see footnote on p. 851) in which the (non-protein) amino-acid *ornithine* acts as "catalyst" (Fig. 568). Ornithine reacts with CO_2 and 1 mol. of NH_3 to give *citrulline*, which itself reacts with a second mol. of NH_3 to give *arginine*; the amidine group of arginine is split off as *urea* under the influence of the hydrolytic enzyme *arginase*, and ornithine is regenerated to continue the cycle. The interaction of amino-acids is further complicated by the fact that the amino-acids glutamic and aspartic acid are involved in the entry of NH_3 into the "cycle." By the use of isotopes, it has been shown that the CO_2 used in urea formation, comes solely and directly from blood HCO_3'.

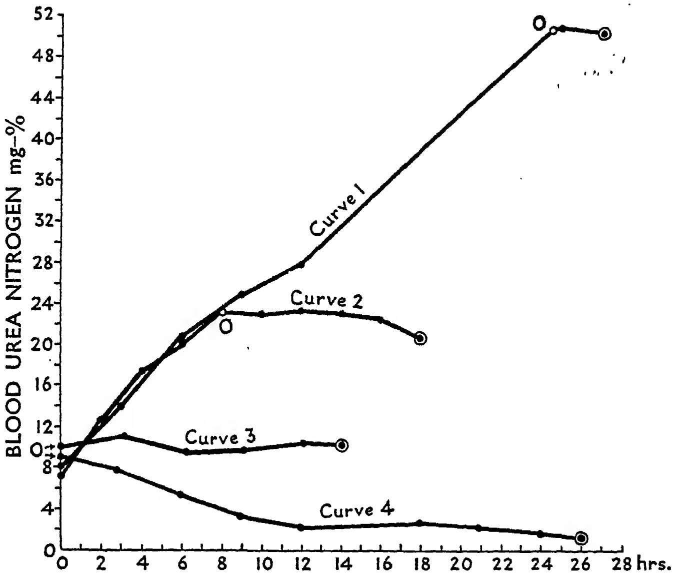


FIG. 567.—Relation of Hepatectomy to Blood Urea Level.

Ordinate: blood urea nitrogen in mg-%.
(1) Kidneys removed; liver removed 24 hours later (at O).
(2) Kidneys removed; liver removed 8 hours later (at O).
(3) Simultaneous removal of liver and kidneys (at O).
(4) Liver removed at O, and urinary secretion maintained (cf. Fig. 566).
(Redrawn from Mann *et al.*, *Amer. J. Physiol.*, 1924, *69*, 382.)

Fate of Non-Nitrogenous Residues from Amino-Acids.—In general, the non-nitrogenous residues remaining after deamination enter the "common metabolic pool" and are either completely dissimilated to CO_2 or (less often) are built up into other body constituents.

Some amino-acids are *glucogenic, i.e.* they may give rise to glucose; some are *ketogenic* and may give rise to acetoacetic acid (and the other "ketone bodies"). The experimental methods employed in the study of these reactions are:

(i) Incubate liver slices with an amino-acid, or perfuse the liver with an amino-acid, and determine whether there is an increase in the glucose or ketone level in the medium or venous outflow.

(ii) Study the diabetic or phloridzinised animal which excretes large amounts of glucose in the urine. The administration of protein, like the administration of carbohydrate, increases the urinary glucose excretion, suggesting that food protein can give rise to glucose. If specific amino-acids are administered, some (*glucogenic*) increase the glucose pool and the degree of glucosuria while others (*ketogenic*) increase the excretion of ketone bodies in the urine. Glucogenic amino-acids cannot counteract hypoglycaemia in hepatectomised animals; thus the liver is an indispensable factor for their transformation into blood glucose.

(iii) Administer amino-acids labelled with C isotopes, and investigate the isotope distribution (if any) in subsequently isolated glucose, glycogen, or ketones.

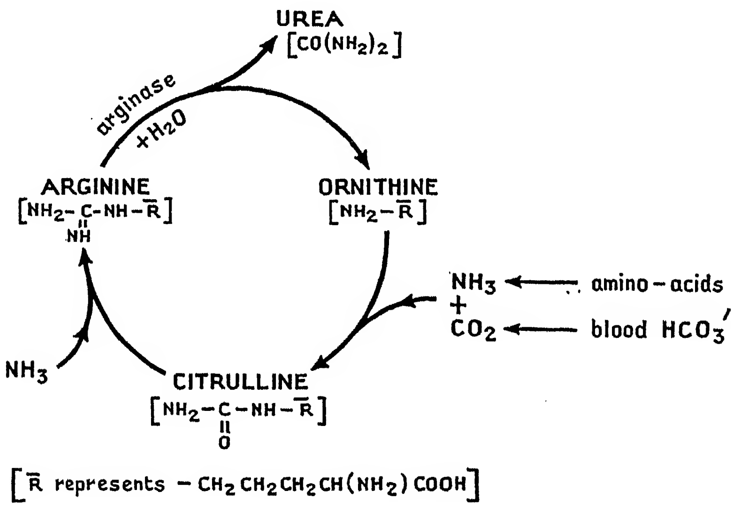


FIG. 568.—Krebs Urea Cycle.

The results of such studies are as follows (Fig. 569):

(*a*) The fate of the non-nitrogenous residues derived from the essential amino-acids, tryptophan, lysine, histidine, and methionine is unknown.

(*b*) Leucine, isoleucine, and phenylalanine (three of the essential amino-acids) and tyrosine are ketogenic. Their deamination and normal oxidation in the liver gives rise to acetoacetic acid, by known irreversible routes; this acetoacetic acid joins the "common metabolic pool" and is subsequently dissimilated (p. 872).

(*c*) *All* the *non-essential* amino-acids are *glucogenic*. This fact indicates that they give rise to an intermediary found on the reversible pathway of carbohydrate dissimilation. Since all the non-essential amino-acids can be synthesised in the body from carbohydrate and a N source (p. 881) the route by which they enter the carbohydrate pathway (including the deamination stage) must be reversible. The identity of the intermediary has been determined in most cases:

(*a*) On oxidative deamination or transamination, alanine and the amino-

acids which can be converted into it (glycine, serine, cysteine) give rise to *pyruvic acid*,

$$\underset{\text{(alanine)}}{CH_3.CH(NH_2)COOH} \xrightleftharpoons{+O,\ -NH_3} \underset{\text{(pyruvic acid)}}{CH_3.CO.COOH}$$

(β) Glutamic acid (and probably proline and arginine) give *a-keto-glutaric acid*,

$$\underset{\text{(glutamic acid)}}{COOH(CH_2)_2.CH(NH_2)COOH} \rightleftharpoons \underset{\text{(keto-glutaric acid)}}{COOH(CH_2)_2.CO.COOH}$$

(γ) aspartic acid gives *oxaloacetic acid*,

$$\underset{\text{(aspartic acid)}}{COOH.CH_2.CH(NH_2)COOH} \rightleftharpoons \underset{\text{(oxaloacetic acid)}}{COOH.CH_2.CO.COOH}$$

These three products are all components of the citric acid cycle (Fig. 569, and p. 852, Fig. 559); they can either be built into glucose, or dissimilated in the liver to yield energy, thus "sparing" the dissimilation of glucose.

It is found by method (iii) *supra* that the carbohydrate apparently formed from glucogenic amino-acids does not always arise by direct conversion of the amino-acid to glucose. It can also come from the conversion to glucose

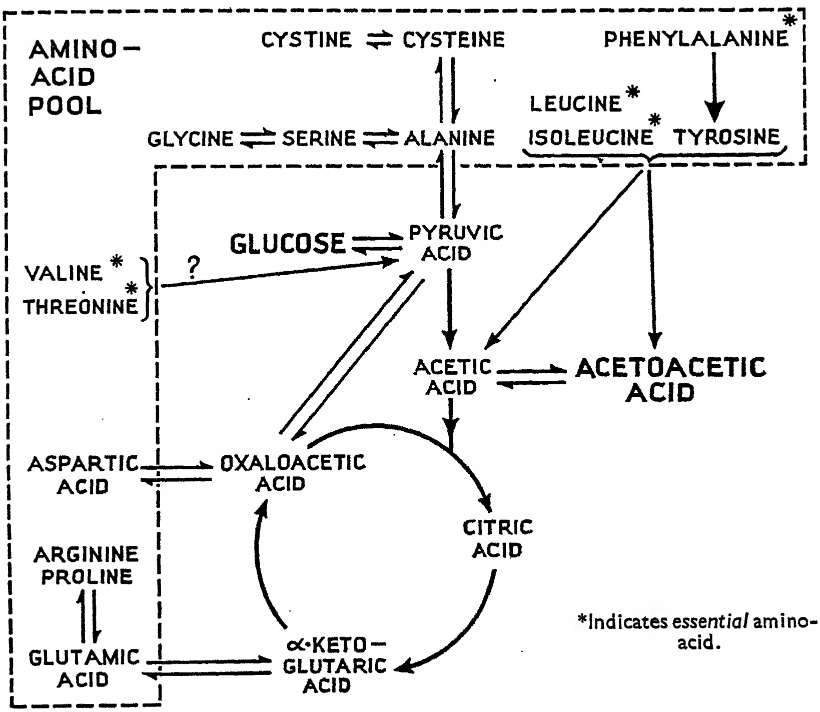


FIG. 569.—Fate and Inter-relationships of some of the Amino-acids. (Diagram by J. B. Jepson.)

of some other metabolite which has itself been " spared " by the non-nitrogenous residue of the amino-acid.

(*d*) The remaining essential amino-acids (threonine and valine) are glucogenic. How they reach the carbohydrate pathway is not known, but presumably it is by a non-reversible route as they are indispensable in the diet.

FATE OF RESIDUES FROM AROMATIC AMINO-ACIDS.—In certain very rare congenital disorders, the normal oxidative mechanisms of phenylalanine and tyrosine metabolism are deranged by " blockage " at different points. This results in the urinary excretion of intermediate metabolites, and has given valuable information on the normal oxidation of these amino-acids.

The following conditions are of particular interest :

(i) *Imbecillitas phenylpyruvica* (Fölling's disease), a form of idiocy strangely coupled with the inability to convert phenylalanine to tyrosine. Phenylalanine accumulates, and phenylpyruvic acid (the deamination product of phenylalanine), appears in the urine.

(ii) *Alcaptonuria*,[1] where *homogentisic acid* (2:5-dihydroxy-phenyl-acetic acid)appears in the urine ; such urine darkens considerably when alkaline and exposed to air, and may turn almost black. Homogentisic acid is apparently a normal metabolite from tyrosine, but alcaptonurics excrete it rather than oxidise it further to acetoacetic acid as is its normal fate (p. 888).

Exogenous and Endogenous Metabolism.—It has been customary to divide the metabolic processes that take place in the body into two fairly sharply defined compartments : (i) metabolism of the *foodstuffs* (*exogenous*) ; (ii) metabolism of the *tissues* (*endogenous*).

The distinction drawn was analogous to that between the consumption of fuel (food) *by* an engine, and the wear and tear *of* the engine itself (tissues). But as has been emphasized, the proteins and nucleic acids and most of the other constituents of the tissue cells are constantly undergoing breakdown processes similar to those that affect the foodstuffs in the gut ; in general, the products of digestion and those of cell disintegration enter a common pool and suffer a common fate (p. 853). There was, however, a fundamental truth in the classical conception : some of the urinary constituents vary widely in amount in relation to the protein-content of the food, while others are relatively independent of the diet. This distinction is brought out in the Table below (from Folin), which shows the differences in the composition of the urine on two diets, both *adequate in calories*, but one *poor* in protein-content.

It is obvious that the total nitrogen excretion in the urine is related to the protein intake. On diets of adequate calorie value but very low in protein, the nitrogenous output may be substantially reduced, *e.g.* from 16·8 to 3·6 g. daily. The excretion of urea and inorganic sulphate is markedly dependent on the diet ; the urinary output of these substances is an index of protein intake, or, in the (discarded) classical terminology, of *exogenous metabolism.* On the other hand, the excretion of creatinine, neutral sulphur and about half the uric acid is quite independent of the protein intake ; in the classical terminology, the output of these substances is an index of *endogenous metabolism.* This term can be usefully retained, since the " continuing metabolism " of the tissues is normally an equilibrium between simultaneous synthesis and breakdown and does not have to be taken into account when making com-

[1] Neuberger *et al.*, *Biochem. J.*, 1947, *41*, 438.

	Protein-Rich Diet.		Protein-Poor Diet.		
Volume of Urine .	1170 c.c.		385 c.c.		Greatly reduced.
Total nitrogen .	16·8 g.		3·6 g.		,,
Urea nitrogen .	14·7 ,,	=87%	2·2 ,,	=62%	,,
Uric acid nitrogen	0·18 ,,		0·09 ,,		Halved.
Ammonia nitrogen	0·49 ,,		0·42 ,,		Unchanged.
Creatinine nitrogen	0·58 ,,		0·6 ,,		,,
Inorganic SO_4 .	3·27 g.	=90%	0·46 g.	=60%	Greatly reduced.
Ethereal SO_4 .	0·19 ,,		0·10 ,,		
Neutral sulphur .	0·18 ,,		0·20 ,,		Unchanged.

parison between dietary intake and urinary output. The end-products of endogenous metabolism appear to represent a slow, continuous, and wasteful " seepage " of useful materials from the common metabolic pools.

Origin and Physiological Significance of Nitrogen-Containing Constituents of Urine.—Mixed proteins contain an average of 16% of N; thus the dissimilation of 100 g. of protein (the recommended, though high, daily intake) yields 16 g. of waste N, nearly all of which is excreted in the urine. The chief N-containing waste products in the urine are: urea; creatinine and creatine; ammonium ions; uric acid. A small quantity of amino-acids is also lost from the body in the urine.

(1) **Urea.**[1]—(i) *Effect of Protein Intake.*—Urea is formed in the liver from ammonia derived from amino-acids (or from ingested ammonium salts); the details of its mode of formation are given on p. 886. As previously explained, the amino-acids of the body constitute a " pool " into which amino-acids pass from the food and the tissues, and from which amino-acids are taken, to be synthesised into protein or otherwise transformed as required by the tissues. Normally, " surplus " amino-acids (whatever their origin) are deaminated and their ammonia is converted into urea; the adult man has no means of storing surplus amino-acids except by a very limited increase in tissue protein. On a normal protein intake the amino-acid pool is " overflowing " with amino-acids which have come predominantly from *recently ingested* food. On such a diet, the urinary urea is therefore mainly of food origin; in other words it is mainly (but by no means exclusively) exogenous. Within limits, *the urea output varies directly with the recen. protein intake.* Urea is always the principal nitrogenous constituent of the urine, constituting 60–90% of the total urinary N. On a normal mixed diet, an adult's daily excretion of urea is 15–40 g. (=7–18 g. urea-N).

(ii) *Effect of Protein-Free Diet which is Adequate in Calories (Energy Content).*—When no protein is eaten, amino-acids are contributed to the " pool " by the breakdown of tissue proteins only. One would perhaps expect that an equal amount of amino-acid would be withdrawn from the " pool " for restitution purposes and that urea formation would therefore cease.

[1] For mechanism of excretion of urea see pp. 28 *et seq.*, 69. Urea clearance and the urea clearance test are described on p. 39.

This, however, is not what happens. Instead, the "pool" is slowly and continuously being depleted of its *essential* amino-acids, mainly through their conversion to high-priority *expendable* body components, *e.g.* creatine (p. 894). The remaining amino-acids not used for these special syntheses cannot be rebuilt into the proteins from which they come, because one or more of the required essential (non-synthesisable) amino-acids needed for the protein synthesis has been utilised elsewhere; as they cannot be stored they are irreversibly broken down, giving rise first to ammonia and then to urea. In other words, more tissue protein is broken down than is rebuilt, the net loss appearing partially as expendable metabolites (*e.g.* creatinine) but mainly as urinary urea. The urinary urea on a protein-free diet is, in the classical terminology, wholly endogenous, and its minimum level is about 4 g. of urea per day. It should be emphasised that it is under the conditions of *this* experiment (*i.e.* protein-free diet containing the full calorific requirements in the form of carbohydrate and fat), rather than in starvation, that the *minimum* excretion of N and of urea takes place.

(iii) *Effect of Starvation (i.e. no protein, no calories).*—In complete starvation, tissue protein (particularly muscle protein) is broken down to amino-acids on a *much larger scale* than in (ii). Some of these amino-acids are used for the restitution of high-priority expendable components as in (ii). But, as no food energy is available, most of the liberated amino-acids are deaminated and the residues utilised for *energy purposes* and to *maintain the blood sugar level*. Tissue protein is thus used in the same way as, and as a substitute for, food protein. In starvation, therefore, urea excretion is on a much larger scale than on a protein-free diet which is adequate in calories, because in the latter condition tissue protein is not being used for "fuel" (cf. p. 900).

(2) **Creatine and Creatinine.**[1]—The structure and relationships of urea, guanidine, creatine, creatine phosphate (phosphocreatine) and creatinine are shown in Fig. 570.

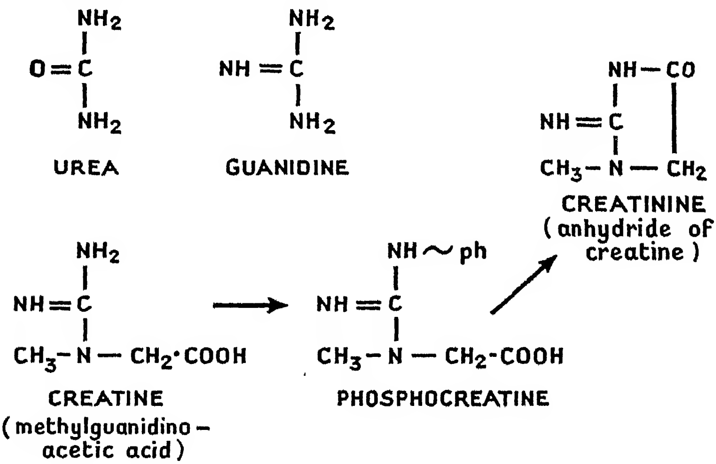


FIG. 570.—The Structures and Relationships of Creatine and Creatinine.

[1] Peters and Van Slyke, *Quantitative Clinical Chemistry*, 2nd edn., vol. I, i, 897.

DISTRIBUTION OF CREATINE.—Creatine, though a constituent of many tissues, occurs in greatest concentration in skeletal muscle, with lesser amounts in heart muscle, brain, and uterus (especially during pregnancy). The creatine concentration of skeletal muscle increases steadily during foetal and post-natal growth until the adult level is finally reached.

In resting muscle, creatine exists largely as *creatine phosphate* (phosphocreatine), the phosphate group of which is linked to the creatine by a high-energy phosphate bond, ~ph (p. 842). Creatine phosphate is formed by the reaction of creatine with adenosine triphosphate (ATP); this reaction is reversible, creatine phosphate transferring its ~ph to adenosine diphosphate, (ADP) and regenerating ATP as required. The energy of carbohydrate dissimilation in muscle is initially made available as the high-energy phosphate bonds of ATP (p. 429), but it is *stored* as the high-energy phosphate bonds of *creatine phosphate*. Creatine phosphate can be accumulated in quantity (unlike ATP), and its high-energy phosphate bonds are available for the rapid re-synthesis from ADP of the ATP which is required for muscular work (Fig. 571).

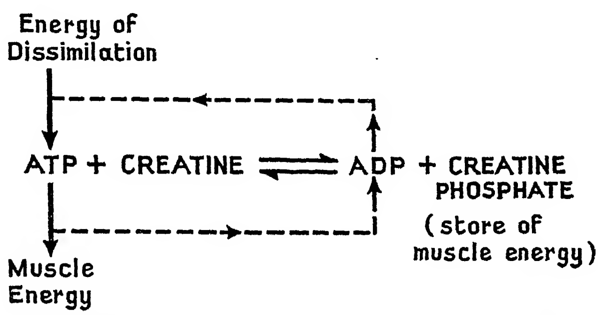


FIG. 571.—Creatine Phosphate as an Energy Store. (Cf. Fig. 259.)

FORMATION OF CREATINE AND CREATININE.—Isotope experiments have shown that creatine is synthesised in the liver from three amino-acids (Fig. 572): *arginine* transfers its amidine group to *glycine* to form guanidino-acetic acid, which is irreversibly transmethylated by *methionine* (p. 883).

The creatine is discharged into the blood (normal level=2–8 mg. per

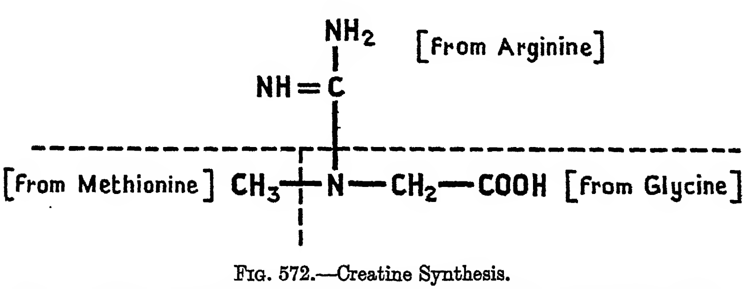


FIG. 572.—Creatine Synthesis.

Note the three amino-acids involved in the formation of creatine, and thence of creatine phosphate and creatinine.

100 c.c. of whole blood, 0·4 mg. per 100 c.c. of serum) and taken up by the muscles as required. Creatine in excess of the muscle storage capacity (*e.g.* after ingestion of large amounts of creatine) is disposed of by unknown means, though in special cases it is excreted in the urine (see below).

The chief reaction of creatine is its phosphorylation by ATP to give creatine phosphate. Some of the store of creatine phosphate is continuously being lost to the body by a slow, spontaneous transformation to *creatinine*, by ring formation with loss of H_3PO_4; the creatinine is excreted in the urine. This is the major source of creatinine, little or none coming *direct* from creatine itself (*i.e.* by ring formation with loss of H_2O).

Excretion of Creatine in Urine (Creatinuria).—Creatine is not a normal component of the urine but may appear in the following conditions:

(i) (*a*) It is constantly and normally present in children of both sexes up to the age of puberty. This fact may be associated with the low storage power of the muscles for creatine at an early age.

(*b*) It is found in an intermittent manner in the urine of women, but is not related to menstruation. There is a continuous creatine excretion during pregnancy. It rises to a maximum of 1·5 g. daily after delivery, when it may be derived from the involuting uterus.

(*c*) Creatine is excreted irregularly by normal men.

(ii) In certain pathological states of muscle, *e.g.* the *myopathies*, creatine is excreted, because of the low storage power of the muscles. Even if only small amounts of creatine are ingested, 90% or more appears unchanged in the urine.

(iii) In any condition in which unusual breakdown of the tissues (especially muscles) occurs, *e.g.* in starvation, diabetes, exophthalmic goitre, and fever (from the increased metabolic rate), creatine is excreted. In all states of under-nutrition the muscles bear the brunt of the burden; as their substance is broken down for energy purposes, creatine is liberated, some of which is stored in the muscles which are still intact, and the rest is passed out in the urine as creatine.

Excretion of Creatinine.—Creatinine is formed in the body exclusively from creatine via creatine phosphate. Minute amounts of creatinine (so-called "endogenous" creatinine) are present in blood (1 mg. per 100 c.c.); it is a non-threshold substance which is excreted in the urine by filtration from the glomeruli (p. 33). If creatinine is ingested, the blood level is raised. This "exogenous" creatinine is also excreted by the tubules (cf. p. 34).

The urinary output of creatinine during a period of work is invariably greater than during the corresponding time of a day of inactivity; this is, however, immediately followed by a period of unusually low output, so that the 24-hour output is unaffected by work. The daily output of creatinine is thus reasonably constant and is independent of the protein intake or the total amount of nitrogen excreted. Creatinine output depends not so much on the muscle mass as on its creatine plus creatine phosphate store, though, of course, the two are related; the amount of creatinine excreted in 24 hours represents the conversion of about 2% of the body creatine.

Urinary creatinine is a product of endogenous metabolism in the classical sense, representing the removal of the steadily expendable creatine. The excretion of creatinine means the loss of valuable methyl groups from the body; of the normal dietary components only methionine can supply the

methyl groups to replace them (p. 882), though tissue synthesis of methyl groups may make some contribution (p. 884).

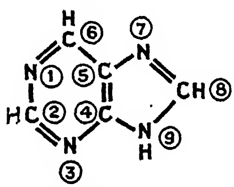
(3) **Ammonium Ions.**—The origin and significance of urinary ammonium ions, NH_4^+, are fully discussed on p. 97.

(4) **Amino-Acids.**—Traces of many amino-acids and small peptides are found in normal urine, representing a mainly "endogenous" excretion. An increase in the excretion of specific amino-acids (probably due to faulty reabsorption by the renal tubules) may occur pathologically (*e.g.* in the Fanconi syndrome).[1] The investigation of urinary amino-acid patterns by "paper chromatography" may have diagnostic value.[2]

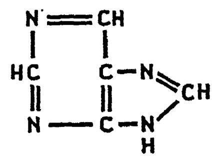
Certain amino-acids are used by the liver to "detoxicate" unwanted compounds (p. 828). These products, of which the chief is *hippuric acid* (from the condensation of benzoic acid with glycine), then appear in the urine.

(5) **Uric Acid.**—Uric acid is the only recognised end-product of purine and nucleoprotein metabolism in man, and is discussed below.

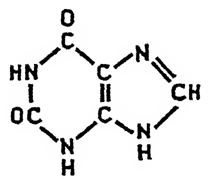
Nucleic Acid Metabolism. Purines and Pyrimidines. Uric Acid.—PURINES.—The formulæ of the chief purines of physiological interest are shown in Fig. 573.



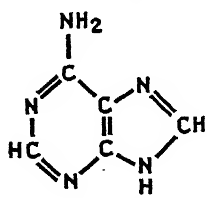
Purine nucleus
showing numbering of different ring atoms.



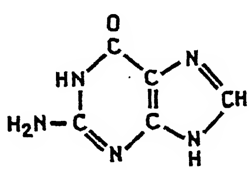
Purine nucleus
older (conventional) representation



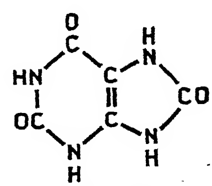
Xanthine
(2 : 6-dioxy-purine)



Adenine
(6-amino-purine)



Guanine
(2-amino-6-oxy-purine)



Uric Acid
(2 : 6 : 8-trioxy-purine)

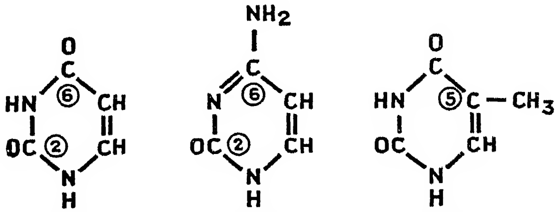
FIG. 573.—The Structures of Some Important Purines.

[1] The Fanconi syndrome is a form of osteomalacia or rickets, resistant to vitamin-*D* therapy, and characterised by urinary loss of phosphate, glucose and amino-acids.

[2] Paper chromatography: a technique for the separation of closely related substances, applicable on the microgram scale. The mixture of substances is adsorbed at one spot on filter-paper; solvents are then passed over the spot by capillary attraction in the paper; each of the components of the mixture forms a separate, discrete spot in a position along the paper dependent on its solubility and strength of adsorption; the spots can be located and identified. For the use of this technique to investigate abnormalities of amino-acid and protein metabolism, see Dent, *Biochem. Soc. Symposium*, No. 3 (1949), 34–49.

Note that purines are built from a system of two *fused* rings each containing two N atoms, namely a 6-membered (pyrimidine) ring and a 5-membered (iminazole) ring. The purines which have been found in nucleic acids are *adenine* and *guanine*.

PYRIMIDINES.—The pyrimidines are built from a 6-membered ring, identical with atoms 1–6 of the purine system. The pyrimidines which have been found in nucleic acids are *uracil, cytosine,* and *thymine*, with the formulæ shown in Fig. 574 :



Uracil (2 : 6-dioxy- -pyrimidine) **Cytosine** (6-amino- -2-oxy-pyrimidine) **Thymine** (5-methyl- uracil)

FIG. 574.—The Structures of Some Pyrimidines found in Nucleic Acids.

DISTRIBUTION OF PURINES AND PYRIMIDINES.—*In Tissues.*—The nuclei of all cells are rich in nucleic acids (present mainly as nucleoprotein, p. 878). All cells contain nucleotides which function as coenzymes (pp. 840, 1028).

Cells contain enzymes which constantly break down tissue nucleic acids and liberate the purines (adenine and guanine) and the pyrimidines which enter a " common purine and pyrimidine pool " ; simultaneously, nucleic acids are rebuilt from the pool. The maturation of nucleated red blood cells to form non-nucleated discs involves the destruction of the nucleus; this process is an important source of purines and pyrimidines to the pool; smaller amounts are provided by the destruction of circulating white blood cells.

In Food.—(i) The cellular foodstuffs which are rich in nuclei, *e.g.* liver, kidney, pancreas, brain, yeast, are likewise rich in nucleic acids which are digested in the intestine (p. 879) giving rise to adenine, guanine, and pyrimidines.

(ii) Oxy-purines, *e.g.* xanthine and hypoxanthine, are present in meat extracts (soup, beef tea, and the like).

(iii) Methyl-purines, *e.g.* theophylline, caffeine, theobromine, are present in tea, coffee, and cocoa.

Purine and Pyrimidine Metabolism.[1]—(1) ADENINE.—There is a " common adenine pool " in the body. The pool receives (i) adenine released from the foodstuffs ; (ii) adenine released from various constituents of the cells (chiefly from nucleic acids, but also from the coenzyme nucleotides and adenosine triphosphate) ; and (iii) adenine synthesised in the body (p. 897). Free adenine is taken from the pool (i) for the re-synthesis of adenine-containing cell components (it is naturally taken up in larger amounts during growth or repair) ; (ii) for conversion into guanine; and (iii) some is "wasted"

[1] Brown, *Cold Spring Harbor Symp. Quant. Biol.*, 1948, XIII, 43.

by being transformed into uric acid (see below). Adenine is not used in the formation of tissue pyrimidines.

(2) OTHER PURINES.—Guanine and other purines (*e.g.* hypoxanthine and xanthine) which have been absorbed from the intestine or formed in the body cannot be reconverted into adenine, nor are they taken up by the cells for synthetic purposes. All these purines are converted into uric acid which is partly excreted in the urine.

(3) FORMATION AND FATE OF URIC ACID.—The stages in uric acid formation in man are probably as shown in Fig. 575 (formulæ in Fig. 573).

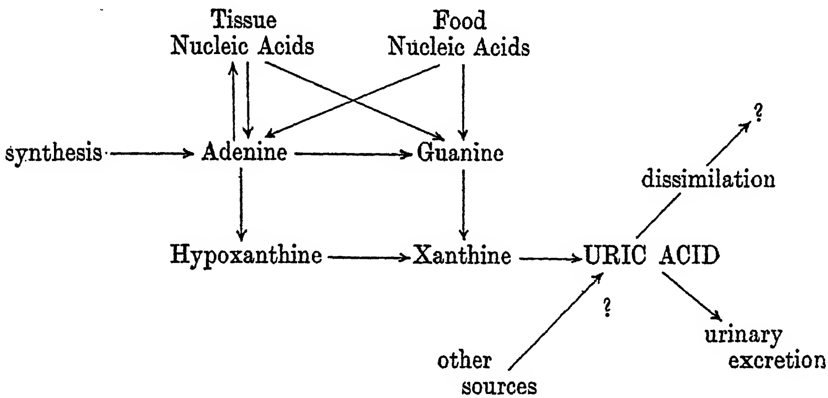


FIG. 575.—The Metabolism of Purines and the Formation and Fate of Uric Acid.

In many animals (*e.g.* most breeds of dog) uric acid is further oxidised to, and excreted as, *allantoin*. Strangely enough, the Dalmatian coach-hound is unable to reabsorb uric acid in the renal tubules and thus excretes all the uric acid which is filtered out in the glomeruli; it is, however, able to form allantoin if the ureters are tied off to retain the uric acid within the body. In man, most of the uric acid formed from purines is excreted in the urine; it has been shown however that a considerable proportion of experimentally administered uric acid is further oxidised to unknown products.

(4) PURINE SYNTHESIS.—Purines can undoubtedly be synthesised in the body. The evidence is as follows:

(i) An infant on a milk diet, which is almost purine-free, puts on weight and increases the total number of its cell nuclei and the purine content of its body.

(ii) A Dalmatian dog fed for a year on a purine- and nucleoprotein-free diet excreted 100 g. of uric acid, of which not more than 10 g. could have come from preformed purine of its tissues.

(iii) There is convincing isotopic-tracer evidence that adenine (and thence uric acid) can be synthesised from *non-purine* components of the common metabolic pool. The *specific* sources of 6 of the 9 atoms of the rings have been determined: C–4, C–5, and N–7 are derived from glycine; C–6 from blood HCO_3'; C–2 and C–8 from one-carbon "formate" units (p. 884).

Guanine is formed from adenine and probably also from non-purine sources.

Arginine and histidine, in spite of resemblances in their structural formulæ to purines are *not* concerned in purine synthesis.

(5) Pyrimidines.—The mode of formation and the fate of the pyrimidines of nucleic acids is still obscure.

Uric Acid.—(1) Formation and Fate of Uric Acid.—See p. 897.

(2) Factors Affecting Urinary Uric Acid Output.—Uric acid is excreted chiefly as monourate salts. The total daily output depends partially on the purine content of the diet, and is 0·75–1 g. on an average. On a purine-free diet it falls to about one-half, *i.e.* 0·3–0·5 g., but this residual (endogenous) excretion varies widely; it is increased by vigorous activity (exercise) and is influenced by the other constituents of the diet.

(i) *Protein* ingestion stimulates the uric acid output: thus changing from a diet of milk and eggs to one of starch and cream lowers the uric acid excretion, though both diets are purine free. Ingested amino-acids, even in small amounts, increase the hourly fasting output of uric acid.

(ii) The *calorific* value of the diet is important. On a constant protein intake, raising the calorific value of the food increases the uric acid output.

(iii) If the diet is mainly carbohydrate with little fat, there is increased uric acid excretion. If the diet consists of little carbohydrate and much fat, the uric acid output is small.

To summarize: uric acid excretion is diminished when the diet has a *low purine* content, a *low protein* content, a *low calorific value*, and consists of *fat* rather than of carbohydrate.

(3) Blood Uric Acid.—The normal limits are 0·7–3·7 mg. per 100 c.c. with an average of 2 mg. It is raised in *gout* up to 9 mg–% (*infra*). The way in which uric acid is carried in the blood is still uncertain. Ingestion of purine-containing food has no effect on the blood uric acid in normal people, but raises it in cases of *renal insufficiency*. It is claimed that uric acid is the first nitrogenous constituent to be retained in nephritis, values of 6 mg–% or over occurring before the other nitrogenous constituents are increased. In *uræmia* (p. 76), values as high as 27 mg–% may be obtained. In *leukæmia* and *pneumonia* considerable breakdown of nuclei of white blood corpuscles occurs, and the blood uric acid is raised. Although in these two diseases the kidneys are excreting a great deal of uric acid they are unable to eliminate it as rapidly as it is formed.

Gout.—Gout is a chronic disorder characterized by (i) excess of uric acid in the blood; (ii) deposition of mono-sodium urate in the cartilages of the joints and in other structures; and (iii) recurring attacks of acute arthritis.

Isotopic tracer work[1] shows that the "uric acid pool" in a normal man is about 1 g., but in a subject with severe gout this was increased to 31 g., a value which included the tophi (solid urate deposits).

The cause of the excessive concentration of uric acid in the blood and body fluids of gouty patients is unknown. There is no evidence of over-production of purine bodies, so that the defect may be one of elimination, or failure of disposal of uric acid in the body. It has been suggested that there is failure of the kidneys to excrete uric acid adequately, although they eliminate the other nitrogenous constituents of the blood, *e.g.* urea, in a normal manner. The uric acid may be circulating in the blood in some

[1] Benedict *et al.*, *J. clin. Investig.*, 1950, *29*, 1104.

abnormal combination, which prevents its filtration by the glomeruli of the kidney.

No satisfactory explanation is available to account for the *deposition of urates* in the cartilages of the joints. The blood uric acid is raised to considerably higher levels in some cases of chronic nephritis and leukæmia without such deposition taking place.

The cause of the *arthritis* in acute gout is also obscure. It is generally assumed that the inflammation is excited mechanically by the deposition of urate crystals in the tissues. Uratic concretions may, however, attain a large size without causing any symptoms whatever.

The excretion of uric acid in the urine is reduced just before an attack of gout and is increased during the attack. Between the attacks it may be about normal or slightly diminished.

The drug *cinchophen* greatly increases the urinary uric acid excretion and lowers the blood uric acid in gout. The increased elimination ceases after 2 days, and if the use of the drug is continued longer it fails to decrease the blood uric acid level further. The excretion of uric acid in the urine is markedly increased by pituitary corticotrophin (ACTH) and cortisone which probably act directly on the kidney, decreasing the degree of reabsorption in the tubules. Acetyl salicylic acid (aspirin) has a similar effect. Colchicine, which may relieve the acute pain of gouty arthritis does not increase the urinary output of uric acid (Fig. 576). The dietary conditions which serve to diminish uric acid excretion have been summarized on p. 898.

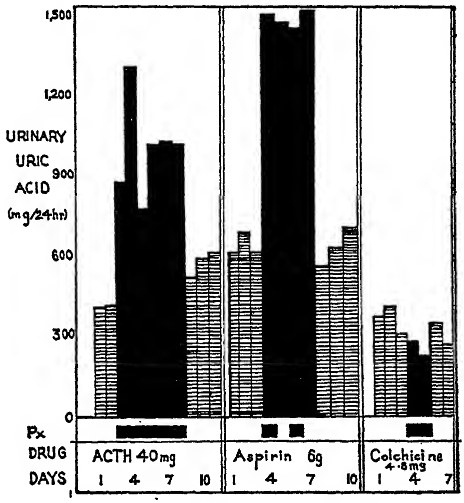


FIG. 576.—Effect of Various Drugs on Uric Acid Excretion. (Benedict *et al.*, *J. Clin. Investig.*, 1950, *29*, 1106.)

Sulphur Compounds in Urine.—The sulphur compounds of urine are derived solely from the sulphur-containing amino-acids (methionine, cysteine and cystine) of the dietary and tissue proteins. The sulphur is excreted in three forms:

(i) *Inorganic Sulphate.*—Sulphur-containing amino-acids of the amino-acid pool that are not used in protein synthesis or other transformations (p. 882) are completely oxidised, the nitrogen appearing as urea and the sulphur as sulphate ions, SO_4''. The sulphate anions are excreted in the urine, with an equivalent amount of cation (Na^+, K^+, NH_4^+). The level of inorganic sulphate excretion follows that of urea excretion, being high on a high-protein diet and reaching a minimum on a protein deficient diet. The normal range of urinary output is 0·3–3·0 g. of sulphate ions per day. Inorganic

sulphate introduced into the body parentally is excreted without appreciable utilisation.

(ii) *Ethereal Sulphate.*—The urine contains small amounts of organic sulphate esters, $R{-}O{-}SO_3H$ (where R=various aromatic radicals), the so-called *ethereal sulphates*. These are the forms in which many phenols, R.OH, are " detoxicated " and excreted in the urine (p. 828). The conjugation of the phenol with sulphur from amino-acids takes place in the liver. The phenols which form ethereal sulphates are: (*a*) physiological compounds (*e.g.* œstrone); (*b*) non-physiological compounds (*e.g.* salicylic acid from ingested aspirin; and (*c*) phenols formed from aromatic amino-acids by the putrefaction of amino-acids in the intestine. Thus tryptophan is decomposed to give indole, which is oxidised in the liver to indoxyl and conjugated with SO_4''. The potassium salt of indoxyl sulphate (called *indican*) is excreted in the urine.

(iii) *Neutral Sulphur.*—This term is applied to a heterogenous collection of unoxidized S-containing substances (*e.g.* cystine, mercaptans) found in the urine in traces; their origins are obscure, but the total neutral S excretion is unaffected by the diet (in the classical terminology, it is endogenous).

Metabolism in Starvation—During starvation, the subject must live on the component tissues of his own body for energy purposes. Even if no physical work is being done, about 2000 Cal. will be needed daily. During the first few days the glycogen stores of the liver are called upon, but these at the best only amount to several hundred grammes and so are of very limited value. The main source of energy must be the fat reserves and tissue protein. So long as fat is available tissue protein is used sparingly, but some *net* destruction of protein cannot be avoided.

Use of Fat.—As discussed on p. 866, the fat which is used in starvation is the neutral fat (triglyceride, *élément variable*) which is found in adipose tissue; it is *mobilized* and taken to the liver, where it is either completely dissimilated or transformed into ketone bodies, which are sent to other tissues for oxidation. The complex lipide (*élément constant*) which forms part of the cell structure is spared till the end.

Use of Tissue Protein.—Tissue protein is treated in starvation like food protein and is hydrolyzed to amino-acids but on a larger scale than normally. The tissues are not treated uniformly: brain and heart only lose 3% of their bulk; muscles, liver, and spleen lose 30, 55, and 70% respectively. There is evidence that the breakdown of tissue protein in starvation is controlled, possibly through the action of the adrenal cortex.

The released amino-acids enter the " common pool." (i) The first call on the pool is to maintain the structure and so preserve the functional efficiency of the *essential* organs (cf. p. 892); this would also include formation of enzymes and those hormones which are proteins or amino-acid derivatives. (ii) The second call on the pool is to preserve the normal blood sugar level without which brain function fails; as the sugar is steadily withdrawn from the blood by all the organs (and not only by the brain), protein used in this way is acting as a source of energy. The conversion of amino-acids to sugar (or the dissimilation of amino-acids *in place of* sugar) occurs in the liver and involves preliminary deamination; the NH_3 thus liberated is excreted as urea.

The urinary urea output in starvation is an index of tissue protein consumption. When the fat stores are exhausted, protein alone is available for energy purposes and death rapidly results.

Nitrogen excretion during the first week of starvation averages about 10 g. daily (which is far higher than the possible minimum (cf. p. 892)); during the second or third week it may fall to a considerably lower value. The changes in the individual nitrogenous and sulphur-containing constituents are essentially *similar to those found on a low protein diet* (see Table, p. 891). Just before death the urinary nitrogen output rises steeply ("premortal rise") owing to the rapid destruction of tissue protein.

SOURCES OF ENERGY.—An examination of the nitrogenous excretion and of the metabolic rate enables an estimate to be made of the relative amounts of fat and protein which are being combusted. Thus, a starving man on the fifth day of his fast, when his energy output was 2000 Cal., excreted 11·4 g. of N. 1 g. of nitrogen in the urine represents the breakdown of 6·25 g. of protein. The urinary nitrogen indicates that $6{\cdot}25 \times 11{\cdot}4 = 71{\cdot}5$ g. of tissue protein were being broken down, which would yield roughly 300 Cal. (1 g. protein=4·1 Cal.). The rest of the energy, *i.e.* 1700 Cal., must have been derived from about 190 g. of fat (1 g. fat=9·3 Cal.). The fasting body was using protein and fat in the proportion by weight of 71·5 : 190, *i.e.* about 1 : 2·5.

KETOSIS.—Owing to lack of carbohydrate, ketogenesis is stimulated and ketones pass into the blood from the liver faster than they are disposed of by the tissues (p. 870). There is thus a ketosis and ketone bodies appear in the urine. The usual steps are taken to compensate for the tendency to acidæmia, *i.e.* buffering by means of bicarbonate, increased pulmonary ventilation, and fall of alveolar CO_2 tension; increased acidity of the urine; increased NH_4^+ excretion (p. 91). As the depot fat is mobilized, there is a slight rise in the blood fat and fatty liver develops.

BLOOD SUGAR.—This is maintained at a steady level almost to the end; it is formed in the liver from amino-acid residues, glycerol (from fats) and lactic acid (from partial dissimilation of muscle glycogen); if hepatectomy is performed, a hypoglycæmia which is rapidly fatal, results.

Death occurs after about 4 weeks, when the body weight is reduced by 50% The temperature only falls just before the end.

The principal metabolic changes in starvation are illustrated in Fig. 577.

For a discussion of the effects of undernutrition see pp. 1045–1050.

Nitrogenous Equilibrium.—When we say that the body is in nitrogenous equilibrium we mean that the nitrogen intake in the food over a given (long) period equals exactly the nitrogen lost in the excreta (mainly in the urine and, to a minor extent, in the fæces) over the same period. Experiments measuring nitrogen intake and excretion under specified conditions are called *nitrogen balance* experiments. A subject in nitrogenous equilibrium is said to be *in nitrogen balance*; by a curious extension of this terminology, a subject whose *intake* of N is *greater* than the output (*e.g.* in growth) is said to have a *positive nitrogen balance*, and one whose *intake* of N is *smaller* than the output (*e.g.* in starvation) is said to have a *negative nitrogen balance.* Equilibrium will be maintained on a diet of protein alone if protein is administered in sufficient quantity to supply the *full energy requirements* of the body. This can readily be done in carnivorous animals, but only with difficulty in man. Let us consider the case of the man mentioned previously whose fasting requirements were 2000 Cal. The ingestion of protein will raise the energy output further and increase his food requirements. Lean meat contains

about 20% of protein; thus 1 lb. (500 g. approximately) yields 400 Cal., so that about 6 lb. would have to be consumed daily to provide the energy needs in this subject at rest, and considerably more would be necessary during exercise. Eskimos consume about 8 lb. meat a day and maintain excellent health, and the same feat has been performed by explorers in Arctic regions.

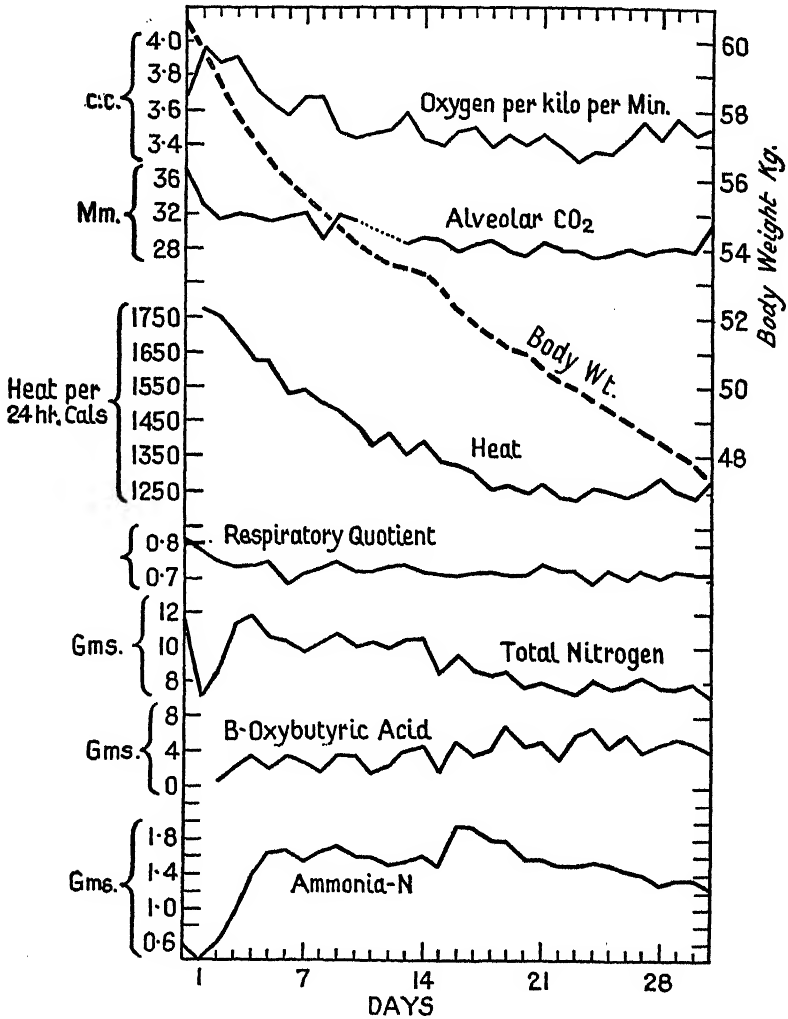


FIG. 577.—Metabolic Changes during 31 days' fasting in Man. (Modified from McLeod, *Physiology and Biochemistry in Modern Medicine.*)

The chart is prepared from an experiment by F. G. Benedict in the Nutrition Laboratory of the Carnegie Institute, Boston, 1912.
Alveolar CO_2 in mm. Hg.; Total nitrogen, β-hydroxy-butyric acid and ammonia-N in the urine in g. in 24 hours.

If an amount of protein is given to a *starving* person equivalent to the amount of tissue protein undergoing destruction, there is an almost equivalent increase in the nitrogenous excretion in the urine. This is a surprising result and not yet explained, but the fact remains that so long as the full energy requirements of the body are not met by ingested food, the tissues continue to be broken down excessively for energy purposes; tissue protein and fat are used in the proportions noted above, and nitrogenous loss continues.

If carbohydrate and fat are added to provide energy, much less food protein is necessary to produce nitrogenous balance. These substances are termed *protein-sparers*. By giving enough carbohydrate to a fasting person it is possible to lower the N excretion to one-third its previous level. Similarly, the daily N excretion of 16 g. on an ordinary diet can be greatly reduced (*e.g.* to 3·6 g. (p. 891)), by giving adequate supplies of carbohydrate and reducing the protein consumption. The *minimum nitrogenous excretion occurs when unlimited amounts of carbohydrate and fat are given to provide the full calorie requirements* (p. 892). But the subject is not in nitrogenous equilibrium; to establish nitrogenous balance certain minimum amounts of protein must be provided in the food. This *irreducible protein minimum* varies with the kind of protein given. The relative values probably depend (i) on the amount of the essential amino-acids contained in these proteins, and (ii) on the degree to which the given protein contains amino-acids in the same proportions as are found in the tissues of the animal in question. It is impossible to maintain N equilibrium on diets which *are deficient in any one* (*or more*) *of*

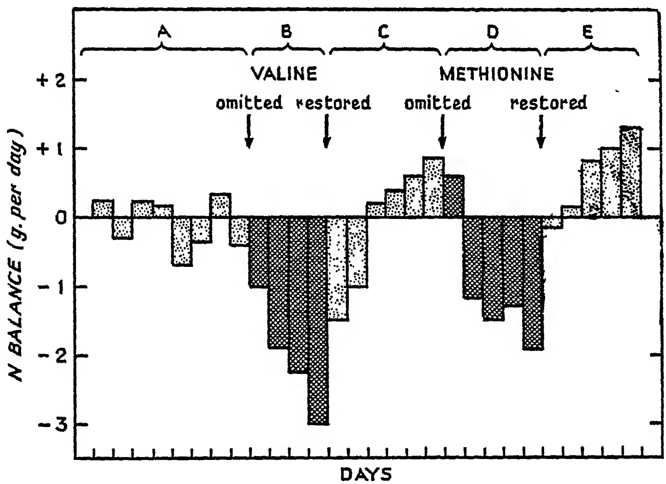


FIG. 578.—The significance of Essential Amino-Acids in the maintenance of Nitrogenous Equilibrium in Man. (Redrawn from Rose, *J. biol. Chem.*, 1950, *182*, 541.)

The subject was on a basal diet, total N content 7·04 g./day, 95% of this N being furnished by a known mixture of 10 pure amino-acids. On this diet, a fluctuating nitrogenous equilibrium is maintained (period A). Omission of *valine* (period B) or *methionine* (period D) immediately causes a negative balance to appear (*e.g.* total N loss over B = 9 g.). Restoration of the missing amino-acid is followed by the establishment of a positive balance (periods C and E).

the essential amino-acids (p. 881), no matter how much protein is consumed (Fig. 578).

100 g. protein are usually recommended in the diet, at least some of it of animal origin. If the right proteins are used nitrogenous equilibrium can be maintained on considerably less, but a bigger protein ration allows a factor of safety should proteins of lower biological value be used, and ensures that adequate amounts of the essential amino-acids which are indispensable for bodily welfare will be ingested (cf. p. 1051). A mixed protein ration drawn from various sources provides a similar safeguard (cf. Fig. 563).

Nitrogen can only be retained in the body as protein, and then only in the following circumstances: (i) during growth; (ii) after recovery from wasting diseases. In the case of a healthy, well-fed, fully grown man, increased ingestion of protein merely results in increased nitrogenous excretion (*i.e.* N-balance is maintained); however, some animals (*e.g.* the cat) may double their muscle bulk on a high protein diet.

THE STUDY OF CELL METABOLISM

The Problems.—In the study of the detailed metabolic processes of living organisms, three main problems arise:

(1) To identify the usual nutrients and metabolites required by the system, be it whole animal or single cell.

(2) To investigate the pathways by which metabolites (or compounds foreign to the systems) are incorporated, utilized and dissimilated by the cells.

(3) To elucidate the function of specific cells, especially in relation to the requirements and functions of the *whole organism*.

The two experimental variables are the organism (or organ, cell, or enzyme) and its metabolite (or substrate): we can study either (*a*) normal systems supplied with metabolites unusual in type or quantity, or (*b*) abnormal systems supplied with usual metabolites. The many different techniques based on these variables are described below, with indications of their applicability to the solution of the three main problems of metabolism.

Experimental Techniques.—(i) To identify the indispensible metabolites and accessory factors required by a system, an *under-feeding* technique is employed, with a basal diet supplemented as desired (*e.g.* investigations on essential amino-acids, pp. 881, 903; vitamins, p. 1019). This is usually the only approach to Problem 1.

(ii) The excreta or tissues of a whole animal are examined chemically for possible intermediates and end-products derived from an initial metabolite, X. Intermediates may be more readily recognized after the administration of unusually large doses of X (the *over-feeding* technique), *e.g.* the feeding of large amounts of creatine followed by measurement of the urinary output of creatine and creatinine (p. 894). Unfortunately, the administration of such large doses may cause unusual side-products to appear which may be mistaken for the normal intermediates. For example, ingestion by a normal animal of large amounts of fatty acids causes acetoacetic acid to appear in the urine (ketosis), but this does not mean that acetoacetic acid is a normal *intermediate* of fatty acid oxidation (p. 871). Suspected intermediates must be administered separately to see if they yield the same final products as X,

and at comparable rates. This technique was formerly a favourite approach to Problem 2 but is now usually superseded by (iv) below.

(iii) A natural metabolite is replaced in the normal system by a *chemically modified* form, whose intermediate and final metabolic products can be more readily recognized. Such modification may involve adding chemical groups amenable to analysis, or adding stabilizing groups which arrest complete breakdown. For example, the first evidence for the breakdown of fatty acids by the stepwise removal of 2C units (p. 871) came from the administration to dogs of phenyl-substituted fatty acids, $C_6H_5.(CH_2)_n.COOH$; acids where n was an even number caused an increase in urinary benzoic acid, $C_6H_5.COOH$, but where n was an odd number there was excretion of phenylacetic acid, $C_6H_5.CH_2.COOH$. There is, however, serious objection to this technique, since even a very minor chemical modification of a metabolite may completely alter its fate in the organism, and may cause derangements in related metabolic systems.

(iv) The use of *isotopes* as atomic "labels" which can be followed throughout their metabolic wanderings has brilliantly overcome the objection to chemical modification mentioned in (iii), since isotopically labelled metabolites are treated in exactly the same way as the usual, unlabelled metabolites. Further, labelled metabolites can be administered in amounts so small as to cause no change in the systems (contrast (ii)). This tracing technique is discussed in detail on pp. 906–8.

(v) The fate of the usual metabolites in whole animals with pathological (usually inherited) *derangements of metabolism* is studied (*e.g.* tyrosine in the condition of alcaptonuria, p. 890). This is a classical use of an abnormal system to give information about Problem 2.

Of greater value for biochemical studies than these "inborn errors of metabolism" are the artificially induced mutations of lower organisms (bacteria, moulds, etc.) caused by irradiation with X-rays; these mutants can be readily isolated and grown, and may show "blockage" at one or more steps of any of their metabolic sequences; by the trial-and-error feeding of possible intermediates, the site of the block and the sequence of the intermediates can be found.

(vi) An animal is *surgically treated* to remove or by-pass a particular organ or tissue, and analyses are made on the blood, urine, etc., of the preparation, to obtain information about the metabolic rôle of that organ or tissue (*e.g.* hepatectomy or the establishment of an Eck's fistula to show the part played by the liver in nitrogenous metabolism (p. 886); pancreatectomy to show the influence of insulin on glucose metabolism (p. 909). Such abnormal systems may also arise through circumstances other than surgical treatment, *e.g.* a diabetic condition may arise spontaneously, or by dosing with *drugs* (*e.g.* alloxan, p. 910). A partial answer is thus provided about the rôle of specific cells of the organism (Problem 3), (but see (vii) below).

(vii) *Isolated* and *perfused organs* are maintained under "physiological" conditions (*i.e.* under conditions which may reasonably be expected to obtain in the whole, unmolested animal) and treated with suitable metabolites (*e.g.* the use of adrenal glands perfused with steroid solutions to study the formation of adrenal corticoids, p. 960). This technique calls for considerable surgical skill, and is unsuited to micro- or multiple-experiments. It has been largely replaced by the use of *tissue slices*, which apparently continue to

function for a limited time in a manner closely resembling the whole tissue (*e.g.* the synthesis by liver slices of cholesterol from acetate (p. 799)). Although it is generally assumed (with justification in many cases) that results from tissue slices hold good for the whole tissue under similar conditions, it must be emphasised that the metabolism of an isolated organ is not necessarily the same as that of the organ in the intact body (consider, for example, the difficulties of interpreting the metabolism of a "normal" eviscerated preparation (p. 911). It is uncertain to what extent it is justifiable to equate the metabolism of a compound administered from *outside* the cell with that of the same compound produced as an intermediate *within* the systems of the cell.

(viii) Enzyme sequences are investigated using finely minced *tissue homogenates*, which contain some or all of the enzymes of the original but no organized cell structures (*e.g.* investigations on the citric acid cycle using liver homogenates to which the usual metabolites are added, p. 851).

Conclusions from the homogenate technique are open to the objection that they are like detailed plans of a house drawn up solely from the pattern of bricks found after the house has been bombed. Nevertheless much of the knowledge of individual intracellular enzymes has been gathered in this way; in the absence of indications to the contrary it is assumed that the normal processes of synthesis and dissimilation that occur within the cells of an organism are the same as those processes found in homogenates of those cells.

The Investigations.—For any particular nutrient or factor (found by (i), and preferably identified), the broad paths of complete metabolism are first mapped out from work with the whole organism (*e.g.* the classic work of Folin on urinary products derived from the N and S of administered protein, p. 890). Then these paths are separated into the sequences and steps of "intermediate metabolism" by the techniques outlined above. It should be the aim to produce an artificial *reconstructed system* using only pure metabolite, purified extracted enzymes and appropriate accessory factors, which will reproduce *in vitro* the overall result or particular sequences of it, implying a relatively complete knowledge of a possible natural metabolic system. Thus the entire process of the anærobic dissimilation of glycogen to lactate (Fig. 556) has been achieved *in vitro* using purified, cell-free enzymes. But it must again be stressed that the demonstration of a transformation in a particular type of preparation cannot be taken as certain evidence that the same transformation occurs with the whole intact organism. Evidence from one technique should always be confirmed by other techniques with inherently different kinds of interference with the normal system.

Isotopes in Metabolism Studies.[1]—A chemical element may exist in several modifications called *isotopes* of that element. Isotopic atoms of a given element differ only in the *mass* of the atomic nucleus. They can be detected and distinguished physically, but *cannot be distinguished chemically.* Most natural isotopes have stable atomic nuclei (*stable isotopes*) and are detected by properties dependent on their different atomic masses (*e.g.* using a mass-spectrometer); the remaining isotopes have unstable, radioactive nuclei (*radio-isotopes*) and can be readily detected and distinguished by their emitted radiations.

[1] Hevesy, *Radioactive Indicators*, N.Y., 1951. Kamen, *Radioactive Tracers in Biology*, N.Y., 1951; *Symposium on Isotopes in Biology and Medicine*, Wisconsin, 1948. Schonheimer, *Dynamic State of Body Constituents*, Harvard, 1942.

Examples.—Natural *carbon* (in the free state, in atmospheric CO_2, and in *all* its compounds) consists mainly of atoms of atomic weight 12 (written ^{12}C) with a *constant* small proportion of a stable isotope ^{13}C; in addition, two radio-active isotopes, $^{11}C^*$ and $^{14}C^*$, with differing types of radiation and different half-lives (*i.e.* the time for the radio-activity to decay to half-strength), can be prepared artificially. *Hydrogen*, as normally found in water and all organic compounds, consists of 99·98% ^1H and 0·02% of stable ^2H (also called deuterium, D); a radio-active $^3H^*$ (tritium, T) can be prepared artificially. Radio-isotopes are now readily available following the discovery of the chain-reacting atomic pile, which can produce almost limitless supplies of $^{14}C^*$, $^{32}P^*$, $^{35}S^*$, $^{131}I^*$ and many other radio-isotopes of the biological elements. Stable isotopes must be obtained by physical separation from natural materials, and large-scale plant is being developed for the supply of deuterium, ^{13}C, ^{15}N, and ^{18}O.

Principles of Tracer Technique.—Since isotopic atoms cannot be distinguished chemically, it is reasonable to suppose that the metabolic pathways of compounds differing only in their isotopic composition are identical. Isotopes are used as a recognizable "label" or "tag" for an element, ion, atomic grouping, or whole molecule. These labelled atoms are always identifiable in a normal atomic environment until they are diluted below the sensitivity of detecting instruments. Thus abnormal isotopic composition can be used to trace the chemical history of an element, usually incorporated in a compound. Certain assumptions are made in using isotopic tracers: (*a*) That isotopic atoms or compounds containing them are *biologically*, as well as chemically, *indistinguishable*. This is only partly true, since their different physical properties (solubilities, diffusion rates, etc.) will influence the rates of their metabolic reactions. However, no *qualitative* differences are found. (*b*) That the *radiation* from radio-isotopes does not itself alter the metabolic systems. Provided only a "tracer" dose of isotope is used, no significant effects are found. In tracer experiments, no derangement whatever need arise in the normal metabolic systems. Admittedly the systems are supplied with an unusual metabolite but (with the two assumptions mentioned above) the differences are such as cannot divert the normal mechanisms (contrast techniques (ii) and (iii) on p. 904–5). The amount of radio-isotope administered is infinitesimal, though it may be incorporated in a larger amount of ordinary material as carrier.

Experimental Procedures.—If the metabolism of an organic compound is to be studied, it must first be synthesized (chemically or biologically) from simple, isotopically enriched inorganic material, incorporating the isotope at a known, stable position of the molecular structure (*e.g.* $^{35}S^*$-labelled methionine (p. 882) prepared chemically from $^{35}S^*O_4''$; $^{14}C^*$-labelled glucose (p. 912) extracted from plants grown in $^{14}C^*O_2$). The choice of isotope is governed by: the atom to be labelled (there are no suitable *radio*-isotopes of N or O); the duration of the experiment compared with the half-life of the (radio)-isotope; radiation hazards; the dilution of isotope involved in the experiment (it must not be more than about 10^3 with stable isotopes, but can be up to 10^9 with radio-isotopes). The simultaneous use of distinguishable isotopes of the same element, or of two labelled atoms in the same compound gives fuller information (*e.g.* the study of multi-labelled glycine with the composition $^{15}NH_2.^{13}CH_2.^{14}C^*OOH$, p. 880).

The labelled compound is then introduced into the system to be studied. Its isotope-containing intermediate and final metabolic products are readily detected and identified, and the position and amount of their labelled atoms can be determined to show their exact relationship to the original metabolite.

In addition to their use in the study of chemical pathways and reactions, labelled atoms and compounds can be used in the following types of investigations :

(i) To follow the *physical* pathways, points of accumulation, and rates of absorption of ions, elements, and compounds (*e.g.* ionic transfer in cells, p. 7 ; absorption of iron, p. 209 ; accumulation of iodide in the thyroid (p. 973)).

(ii) To determine volumes and specific compounds by a *dilution* technique, based on homogeneous mixing with chemically identical non-isotopic material (*e.g.* red cell volume, p. 10 ; body fluid spaces, p. 11).

(iii) To identify and isolate compounds on a microscale (*carrier technique*).

Results of Isotopic Tracer Studies.—Many of the details of the metabolism of carbohydrates, fats and proteins described in this book have been obtained, or positively confirmed, by isotopic tracer techniques. The following are important examples : breakdown and synthesis of fatty acids (pp. 870, 873) ; rate of tissue dissimilation of glucose (p. 912) ; transformations of sulphur-containing amino-acids (p. 882) ; transmethylation reactions (p. 883) ; urea synthesis (p. 887).

In recent years, three new concepts of metabolic processes have been formulated, and it is difficult to see how these concepts could have developed before the advent of isotopic tracers :

(*a*) The majority of the components of a living system are in a state of *dynamic equilibrium,* in a state of balance between breakdown and resynthesis (*e.g.* protein turnover, p. 879 ; ionic transport, p. 7).

(*b*) The products derived from tissue breakdown are chemically and metabolically indistinguishable from many of the materials obtained directly from the external environment, and the classical distinction between endogenous (tissue) and exogenous (food) metabolism disappears (p. 890). Materials from both sources are equally available, and constitute general or specific *pools* of small molecules representing the magnitude of this availability (*e.g.* final common metabolic pathway, p. 853 ; amino-acid pool, p. 880 ; uric acid pool, p. 898).

(*c*) The complex materials of a living system are built from the small, *specific precursor molecules* of the metabolic pools (*e.g.* purines from glycine, p. 897 ; cholesterol from acetate, p. 799 ; adrenaline from tyrosine and methionine, p. 729 ; creatine from glycine, arginine, and methionine, p. 893).

Radio-isotopes are also used *therapeutically*[1] as sources of radiation which will be localized within the body (*e.g.* $^{131}I^*$ for the treatment of thyrotoxicosis, p. 997 ; $^{32}P^*$ for polycythemia, p. 168).

[1] Low-Beer, *Clinical Uses of Radioactive Isotopes,* Illinois, 1950. Roberts, *Archives Middlesex Hospital,* 1951, *1,* 219.

IX

THE DUCTLESS GLANDS REGULATING METABOLISM[1]

The ductless glands to be considered in this section are :

(1) ISLETS OF LANGERHANS.

(2) ANTERIOR PITUITARY: the hormones concerned with growth and metabolism. (The hormones regulating reproductive activities are discussed on p. 1083, 1113. The posterior pituitary (neurohypophysis) is dealt with on p. 41).

(3) ADRENAL CORTEX.

(4) THYROID.

(5) PARATHYROID.

(6) THYMUS : very little is known about the thymus and it is discussed here for convenience.

ISLETS OF LANGERHANS. INSULIN. DIABETES MELLITUS.[2]

Islets of Langerhans.[3]—Scattered between the ordinary alveoli of the pancreas are small clumps of epithelial cells which are free from ducts, and are supplied by large convoluted capillary vessels (Fig. 579). These *islets of Langerhans* contain two kinds of cells distinguishable by the character of their granules, which can be demonstrated by intravitam staining with neutral red : α granules which are fixed in alcoholic solution ; β granules which are fixed in watery solutions. The nerve supply of the islets is from the vagus (p. 917).

The islets (probably the β cells) secrete the hormone *insulin.* The evidence is as follows :

(i) *Excision of the pancreas* causes a severe and rapidly fatal condition closely resembling the disease diabetes mellitus which occurs in man. The operation can be carried out in two stages : (*a*) the greater part of the gland is removed and a part of the tail is transplanted under the skin. No glycosuria develops, and rapid recovery occurs (as the pancreatic remnant can form enough insulin to maintain health). (*b*) The transplant is removed later by a simple skin incision. In a few hours the urine contains 5–10% of glucose, and the animal usually dies in coma within three weeks.

(ii) If the *pancreatic ducts are ligated* or blocked, the externally secreting tissue disappears, and the pancreas is converted into a thin strip of connective tissue which still contains the islets. Carbohydrate metabolism at this

[1] Selye, *Text Book of Endocrinology,* Acta Endocrinologica, Montreal, 2nd end. 1950. Pincus and Thimann, *The Hormones,* N.Y., vol. *1,* 1948 ; vol. *2,* 1950.

[2] Soskin and Levine, *Carbohydrate Metabolism,* Chicago, 1946. Peters and Van Slyke, *Quantitative Clinical Chemistry, Interpretations,* vol. *1,* pt. i, 2nd. edn., 1946. Himsworth, *Lancet,* 1939, ii, 1 *et seq.*

[3] Gomori, *Bull. N.Y. Acad. Med..* 1945, *21,* 99.

stage is normal. Removal of this strip, however, leads to the development of intense diabetes mellitus.

(iii) The administration of *alloxan*[1] to animals specifically poisons and destroys the β cells of the islets of Langerhans and produces symptoms of diabetes mellitus.

(iv) The active principle *insulin* can be extracted from the pancreas and also from pure islet tissue.

(v) In *clinical* diabetes mellitus, lesions of the β cells of the islets are sometimes found, but not regularly. The significance of these results is considered on p. 918.

Function of a Cells.[2]—Their function is unknown. A blood glucose raising

Fig. 579.—Structure of Human Pancreas showing Contrast between Islet of Langerhans and surrounding Acinous Tissue (×470). (Maximow and Bloom, *Text Book of Histology,* W. B. Saunders & Co.)

substance, which is alleged to be derived from the α cells, has been extracted from the pancreas. This so-called hyperglycæmia agent is found in minute amounts associated with insulin preparations. The physiological rôle, if any, of this agent, requires further study (cf. p. 918, footnote 1).

Insulin.[3]—Insulin, the internal secretion of the islets, is a soluble protein, with a molecular weight in solution of 48,000; it has been isolated in pure crystalline form and considerable progress has been made with determining the arrangement of the amino-acids along the peptide chains. It is destroyed by proteolytic enzymes in the bowel and therefore has always to be administered parenterally, usually by subcutaneous injection. Like all the hormones which regulate metabolism, insulin probably acts on certain of the tissue enzymes which influence chemical transformations.

Action of Insulin.—Glucose diffuses readily through the capillary

[1] Lukens, *Physiol. Rev.*, 1948, *28*, 304.
[2] Bruch, *Arch. int. Med.*, 1950, *86*, 427. Pincus, *J. clin. Endocrin.*, 1950, *10*, 556.
[3] Jensen, in *The Hormones*, N.Y., 1948, *1*, 301.

endothelia and cell membranes and is thus distributed fairly uniformly throughout the body fluids, both extracellular (including the blood) and intracellular. It should be emphasized that the blood glucose level is an approximate guide to the concentration of glucose in the body fluids. The total body-fluid glucose content obviously depends on the rate at which glucose enters and leaves these fluids. Insulin decreases the glucose concentration in the body fluids (and thus lowers the blood glucose level) because (i) it increases the withdrawal of glucose from these fluids, and (ii) it decreases the rate of addition of glucose to these fluids.

(1) Insulin *increases the withdrawal of glucose* from the body fluids in three ways:

(i) It increases the deposition of glycogen in the liver and in the muscles (and perhaps also in other tissues).

(ii) It increases the rate of complete dissimilation (oxidation) of glucose to CO_2 in the tissues.

(iii) It increases the rate of conversion of glucose in the liver into fatty acids. It also facilitates the deposition of fat (*lipogenesis*) in adipose tissue from glucose.

(2) It *decreases the rate of addition of glucose* to the body fluids in two ways:

(i) It depresses neoglucogenesis (new glucose formation from the non-nitrogenous residues of amino-acids.)

(ii) It may perhaps depress the rate of conversion of liver glycogen into glucose.

Methods of Study.—The action of insulin can be studied on many preparations and in many ways:

(i) In the whole animal, either normal or depancreatized or subjected to other operative procedures.

(ii) In the eviscerate preparation. By evisceration is meant removal of the liver, kidneys and intestine; under these conditions there is no absorption of monosaccharides from the intestine and none can be lost in the urine. The liver, which is the prime regulator of the blood glucose level is excluded. (If the brain is removed as well, convulsions due to hypoglycæmia (p. 915) are prevented). The preparation thus consists essentially of a circulation, artificially ventilated lungs, and muscles. If the *pancreas is left intact*, the preparation is called (by an illuminating convention) "normal"; if the *pancreas also is removed*, the preparation is described as depancreatized or "*diabetic*".

Soskin has used this preparation to determine the rate of what he calls glucose "utilization", *i.e.* the amount of glucose which is withdrawn from the body fluids and either completely dissimilated (oxidized) to CO_2 or transformed into other substances, but excluding the glucose which is converted into glycogen, other hexoses or lactic acid.

(iii) On perfused isolated organs and tissue slices.

The use of isotopically labelled substances in any of the above preparations greatly simplifies the interpretation of the results.

Some of the actions of insulin are considered below in greater detail.

1. Effect of Insulin on Rate of Glycogen Deposition.—(1) Glycogen can be formed from glucose by isolated tissues (*e.g.* muscle *in vitro*) in the absence of insulin if the appropriate enzyme systems are present. The rate of glycogen formation increases as the concentration of glucose in the medium

is increased. Thus Fig. 580, lower curve (insulin absent), shows that when the glucose concentration in the medium was increased from 100 to 400 mg–% the amount of glycogen formed by the muscle rose from 150 to 350 mg./100 g. Similarly in the intact animal, the rate of glycogen deposition in the muscles and in the liver is directly related to the blood glucose level.

(2) *In vitro*, the *addition of insulin* to the medium increases the rate of glycogen deposition, particularly when the glucose concentration in the medium is within physiological limits. A comparison of the lower and upper curves in Fig. 580 shows that when the glucose concentration in the medium was 100 mg–% the addition of insulin increased the rate of glycogen deposition about threefold.

(3) Similar results are obtained *in vivo*, *e.g.* in eviscerate preparations both depancreatized and " normal ".

2. **Effect of Insulin on Rate of Glucose Dissimilation and " Utilization ".**[1] — (1) STUDIES OF GLUCOSE " UTILIZATION." — The rate of glucose " utilization " (in Soskin's sense, p. 911) is directly related to the blood glucose level. In the " normal " eviscerate preparation an increase in the blood glucose level (within certain limits) increases the rate of glucose " utilization ". In the depancreatized preparation a similar relationship holds good, but the rate of glucose " utilization" at any given blood glucose level is considerably lower than in the " normal " preparation. These findings are illustrated by the two curves in Fig. 581. Thus at a blood glucose level of 100 mg–% the " normal " preparation " utilized " 250mg. kg. hour compared with only 100 mg. in the depancreatized preparation. These observations show that insulin increases the " utilization " of glucose by the tissues.

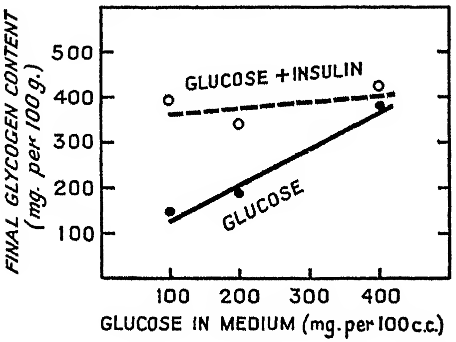


FIG. 580.—Influence of Glucose Concentration on Deposition of Glycogen in Muscle (rat diaphragm) *in vitro*, in the Presence or Absence of Added Insulin. (Soskin *et al.*, *Proc. Soc. exp. Biol. Med.*, 1941, *46*, 390.)

(2) STUDIES OF GLUCOSE DISSIMILATION.[2]—Glucose containing $^{14}C^*$ is injected into an animal, and the $^{14}C^*$ concentration in the blood glucose and in the CO_2 of the expired air is determined. From these data and the volume of total body water the amount of glucose which has been completely dissimilated (oxidized) to CO_2 in a given time can be determined. Such studies prove that glucose can still be completely dissimilated in the absence of insulin. The depressant effect of insulin lack on the rate of glucose dissimilation varies with the species. Thus in the rat made diabetic with phloridzin the rate of glucose utilization is 85% of the normal. In the completely depancreatized dog, however, the rate falls to 40% of normal. It should be remembered that

[1] When " utilization " is written thus (in inverted commas) it is being used in Soskin's sense.

[2] Chaikoff *et al.*, *J. biol. Chem.*, 1950, *187*, 571; 1951, *188*, 865.

these diabetic animals are markedly hyperglycæmic and that their high blood glucose level compensates in part for the absence of insulin, by directly increasing the rate of glucose dissimilation. The injection of insulin into these animals increased the rate of glucose dissimilation in spite of the fact that the blood glucose level was simultaneously reduced. The effect of insulin lack and of the blood glucose level on the rate of glucose dissimilation in man has not yet been examined by tracer methods.

3. Studies of Conversion of Glucose to Fatty Acids in the Liver.—Studies on liver slices using tracer methods ($^{14}C^*$ glucose) have shown that in normal animals on a high carbohydrate diet the liver rapidly converts glucose into long chain fatty acids, in the main acids containing 16 or 18 carbon atoms. If insulin is added to the medium the rate of this conversion is further increased. If liver slices from diabetic animals are used, it is found that (owing to the absence of contained insulin), glucose can no longer be converted into fatty acids.

4. Studies on Lipogenesis.—In the diabetic, the deposition of fat (in adipose tissue) from dietary glucose is depressed. Thus in the normal rat on a daily glucose intake of 15 g., 5 g. was converted daily into fat; in the diabetic rat the amount of glucose converted into fat was reduced to one-twentieth of normal.[1] If insulin is injected directly into adipose tissue the local deposition of fat is increased.

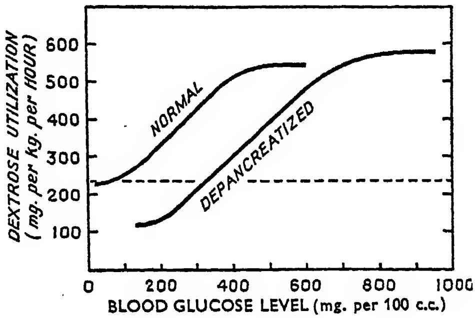


Fig. 581.—Relationship between Blood Sugar Level and Rate of Glucose "Utilization" in Eviscerate Preparations in "Normal" and Depancreatized Dogs. (Soskin and Levine, *Amer. J. Physiol.*, 1937, *120*, 761.)

5. Studies on Neoglucogenesis.—The evidence that insulin inhibits hepatic neoglucogenesis is given on pp. 857 *et seq.*, 914, 920.

Relation of Insulin to Tissue Enzyme Systems and to other Hormones.—Before glucose can "enter the metabolic machinery of the cell" it must first be phosphorylated by ATP (adenosine triphosphate) in the presence of the enzyme hexokinase to form glucose-6-phosphate; subsequently the hexose is either built up into glycogen or dissimilated (p. 844). In the absence of insulin the initial phosphorylation reaction does not proceed at the normal rate. Several attempts have been made to explain how insulin promotes the phosphorylation of glucose.

(i) According to Cori, insulin acts indirectly on the enzyme hexokinase. He concludes from experiments on tissue extracts that hexokinase is directly inhibited by anterior pituitary and adrenal cortex hormones which thus depress the rate of glucose uptake by the cells. Insulin (according to Cori) has no direct action on hexokinase but it annuls the inhibitory action of the other hormones; it thus releases hexokinase which can accelerate the phosphorylation of glucose.

[1] Stetten, *Recent Progress Hormone Research*, 1949, *4*, 189.

(ii) But studies on purified hexokinase preparations using pure anterior pituitary hormones, pure adrenal corticoids, and insulin have not confirmed Cori's observations which were made on cruder materials. There is thus no decisive evidence at present that any of these hormones influence the hexokinase system. It is probable that these hormones act on the complex enzyme systems which control the reaction

Free Glucose $\rightarrow$ Hexose phosphate

but the *precise* enzyme which they act on is as yet undetermined. *Insulin promotes this reaction ; the anterior pituitary and adrenal cortex hormones block it.* In the normal animal the actions of insulin and its antagonists are correctly *balanced*. An excess or a deficit of either group of hormones disturbs the rate at which the phosphorylation of glucose proceeds. Thus lack of insulin decreases the rate of glucose uptake by the cells ; excess of insulin has the reverse effect. *Excess* of anterior pituitary or adrenal cortex hormones has the same effect as *lack* of insulin.

(iii) Insulin produces its characteristic effects in the hypophysectomized or adrenalectomized animal ; in fact under these conditions insulin sensitivity is markedly enhanced (p. 926). These observations demonstrate that insulin does not act solely by modifying the action of other hormones.

(iv) Insulin influences other reactions besides the phosphorylation of glucose : its detailed mode of action is unknown. Thus, as already stated, it inhibits hepatic neoglucogenesis and promotes lipogenesis. The former action is antagonized by anterior pituitary and adrenal cortex hormones.

Action of Insulin in the Intact Animal or Man.—(1) HYPOGLYCÆMIA.—As a normal animal is not in need of extra insulin the result of the insulin injection is a pathological *hyper*insulinæmia. The outstanding effect of insulin is to reduce the glucose content of the body fluids ; there is consequently a rapid fall of the blood glucose (*hypoglycæmia*). The symptoms of hypoglycæmia are almost entirely referable to changes in the central nervous system, because circulating blood glucose is the sole source of energy of the brain ; a fall of blood glucose deranges cerebral activity. Hypoglycæmia is a harmful and dangerous state ; the body reacts to it in a manner which tends to restore the blood glucose level to normal. The effects of insulin injection in the normal animal are therefore the algebraic sum of (i) the *direct* action of insulin and (ii) the *compensatory* reactions which occur to overcome the hypoglycæmia.

(2) CAUSES OF INSULIN HYPOGLYCÆMIA.—The fall of blood sugar is due (i) to increased withdrawal of glucose from the body fluids to form glycogen, fatty acids, and fat, and (ii) to decreased flow of glucose into the body fluids. In the normal animal when there is a simple withdrawal of glucose from the body fluids, *e.g.* in exercise, *no* significant hypoglycæmia occurs even in the fasting state because a compensatory increased inflow of glucose takes place. The liver is the site of this reaction : it converts glycogen stores into blood glucose and forms new glucose from amino-acid residues. Insulin *inhibits these reactions of the liver* and so the hypoglycæmia is initially progressive.

(3) RECOVERY FROM THE HYPOGLYCÆMIA.—In the intact animal the hypoglycæmia sets up a series of complex reactions which gradually restore the blood glucose level to normal. The main reactions are as follows :

(i) The hypothalamus is stimulated causing increased secretion of

hormones which antagonize the action of insulin on the liver and tissues; there is evidence for example that the secretion of ACTH is increased which in turn leads to a discharge of adrenal corticoids.

(ii) There is increased sympathetic activity and increased secretion of adrenaline; conversion of liver glycogen to glucose is thus accelerated.

As the concentration of insulin decreases and that of its antagonists increases the withdrawal of glucose from the body fluids falls and the inflow of new glucose rises; ultimately the normal blood glucose level is restored.

The importance of the rôle of *liver* in the compensatory reactions can easily be demonstrated. In the liverless animal, the blood glucose is lowered by insulin but *never recovers*; if insulin is injected into an animal in which the liver has been partially extirpated or damaged, recovery of the blood glucose is delayed and is not complete (Fig. 582).

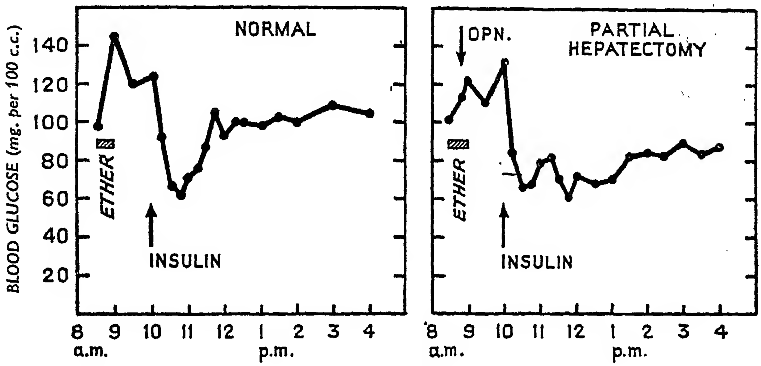


FIG. 582.—Action of Insulin on Blood Glucose before and after Partial Hepatectomy in the Dog. (After Mann *et al.*, *Amer. J. Physiol.*, 1923, *65*, 416.)

Ordinate: Blood glucose in mg./100c.c.
Abscissa: Time in hours.
Normal: ether was given to make conditions comparable with those in the operated animal. Blood glucose rises. At arrow, inject insulin (0·16 unit/kg.). Note rapid fall and rapid recovery of the blood glucose.
Partial Hepatectomy: give ether and remove 75% of the liver tissue. Repeat insulin injection. There is the same fall of blood glucose but recovery is slow and incomplete.

(4) SYMPTOMS OF HYPOGLYCÆMIA IN MAN.—As explained above the principal manifestations of clinical hypoglycæmia are cerebral in origin.

(i) There is a feeling of extreme hunger and a great sense of fatigue; walking becomes difficult. The patient may become anxious, worried, excitable, or behave as if he were intoxicated with alcohol or demented.

(ii) Tremulousness develops, and fine movements cannot be carried out.

(iii) Vasomotor disturbances of cerebral origin occur, *e.g.* flushing and profuse perspiration which may soak the bedclothes, or pallor and a sense of chilliness. The latter symptoms may be due to the compensatory secretion of adrenaline which occurs.

(iv) Later, there are graver mental disturbances, delirium, and convulsions [1]; coma then develops with loss of the deep reflexes.

[1] Insulin hypoglycæmia is sometimes deliberately produced in patients with *schizophrenia* in order to produce convulsions which are alleged to improve the clinical state.

The first symptoms may set in at a blood glucose level of 75 mg–%, or not till a much lower level, *e.g.* 30 mg–% is reached. In chronic diabetics with high blood glucose, symptoms may occur when the hyperglycæmia is suddenly reduced though the blood glucose level is still above normal, *e.g.* at a level as high as 150 mg–%.[1]

(5) Relief of Hypoglycæmia.—The symptoms of hypoglycæmia are relieved by the administration of *glucose*. The other hexose monosaccharides, *e.g.* mannose or fructose (lævulose), are relatively ineffective. *Adrenaline* helps by mobilizing the liver glycogen; this action on the liver is, however, more marked in a well-fed animal than in a starved or diabetic subject, in whom the glycogen stores of the liver are low.

Clinical Hyperinsulinism.—This condition usually occurs in association with hyperfunctioning tumours of the *islets* of the pancreas. The symptoms are those of typical hypoglycæmia; in one case[2] the patient first

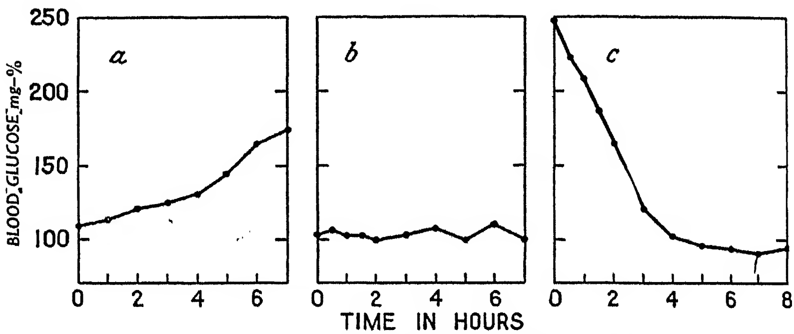


Fig. 583.—Relation of Pancreas to Blood Glucose Regulation.
Blood glucose curves in dogs.

(*a*) Pancreatectomy (progressive rise of blood glucose).
(*b*) Pancreatectomy combined with simultaneous pancreatic graft (normal blood glucose).
(*c*) Previous pancreatectomy leading to rise of blood glucose: pancreatic graft (at O) restores raised blood glucose to normal. (Houssay, *Amer. J. med. Sci.*, 1937, *193*.)

complained of weariness and confusion, and ultimately collapsed towards the end of each day; he said that he felt better after food. As the disease progressed, symptoms developed three hours after meals when the blood glucose was 55 mg–%; four hours after food, perspiration, muscular twitching, and incoherent speech were noted when the blood glucose was 36 mg–%; when it fell to 27 mg–%, stupor set in, but ten minutes after administering glucose, recovery was well established. At operation a pancreatic tumour was found (which could not be removed), consisting of cells resembling islet tissue, and a secondary deposit of similar appearance was seen in the liver. After the operation the condition became worse, and no less than 1000 g. of carbohydrate had to be given daily to ward off hypoglycæmic manifestations.

Control of Insulin Secretion.[3]—(1) Insulin is *continuously* secreted in the normal animal. This is demonstrated by the effects of pancreatectomy; after this operation the blood glucose commences to rise within 2–4 hours

[1] *The insulin test of gastric secretory activity* is discussed on p. 785 (cf. Figs. 515, 517).
[2] Wilder *et al.*, *J. Amer. med. Assoc.*, 1927, *89*, 348.
[3] MacLeod, *Lancet*, 1930, ii, 512.

(Fig. 583, *a*), continues to rise for 1–2 days, and then becomes steady at the new high level.

(2) Insulin secretion can occur in the absence of the pancreatic nerve supply. Thus a successful pancreatic graft (attached to the circulation of the neck of the recipient via the jugular vein and carotid artery) prevents the onset of hyperglycæmia if introduced immediately after pancreatectomy; it restores the blood glucose to normal in an operated animal with fully developed diabetic symptoms (Fig. 583, *b*, *c*); such a graft is, of course, completely denervated. Fundamentally, then, the regulation of carbohydrate metabolism is not dependent on the integrity of the nerve supply to the islets.

(3) The rate of insulin secretion is regulated by the *concentration of glucose in the arterial blood reaching the pancreas.* This, too, can be well demonstrated on a pancreatic graft. If a 2·5% solution of glucose is injected into its artery, there is a transient *fall* in the general blood glucose level, indicative of an increased outpouring of insulin by the islet cells (Fig. 584).

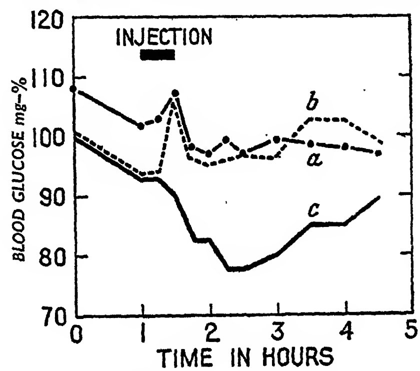


FIG. 584.—Regulation of Insulin Secretion.

Blood glucose curves in pancreatectomized dogs with functioning pancreatic graft. During the period marked INJECTION, fluids were injected as follows:
(*a*) 2·5% glucose into jugular vein (blood glucose unchanged).
(*b*) Locke's solution into pancreatic artery (blood glucose unchanged).
(*c*) 2·5% glucose into artery supplying pancreatic graft. Note marked fall of blood glucose which is evidence of insulin secretion. (Houssay, *Amer. J. med. Sci.*, 1937, *193*.)

(4) The islets, however, receive a rich innervation from the *vagi* which may constitute a "fine adjustment" of secretion of insulin. Stimulation of the vagi is said to lower the blood glucose owing to an outpouring of insulin. Cross-circulation experiments indicate that changes in the blood glucose act on the *central nervous system* (hypothalamus, pons and medulla) and modify insulin secretion appropriately via the vagi (cf. p. 785). After section of the vagi, the blood glucose level is maintained within normal limits, but the glucose tolerance curve declines more slowly than normal. Though the islets are thus partly under nervous control, their secretory activity is mainly regulated by the level of the blood glucose.

(5) The action of *anterior pituitary* extracts and *thyroid* on the islets is considered below, p. 918.

Experimental Diabetes Mellitus.—Diabetes mellitus must be regarded not as a disease but as a *syndrome* (*i.e.* a distinctive group of symptoms and signs), the only indispensable constituents of which are *persistent hyperglycæmia* and *glycosuria.*

Diabetes mellitus as thus defined can be experimentally produced in a number of ways.

(1) By PANCREATECTOMY in animals (p. 909) or man. Recently the pancreas has been completely removed in *man* in cases of malignant disease of the gland. These patients reveal for the first time the effects of

uncomplicated deficiency of insulin in the human subject. The findings closely resemble those found in experimental pancreatectomy or in clinical cases of diabetes mellitus, but the amount of insulin required to maintain a normal blood glucose level is moderate, *i.e.* 25–40 units per day on a diet of 250 g. of carbohydrate, 100 g. of protein, and 100 g. of fat.[1]

(2) THYROID AND HYPOPHYSEAL DIABETES.—Administration of dry thyroid or injection of diabetogenic anterior pituitary extracts induces hyperglycæmia and glycosuria for as long as the treatment is kept up but for no longer. These hormones antagonize the action of insulin.

(3) METAHYPOPHYSEAL DIABETES.—Certain crude anterior pituitary extracts are injected daily, the dose being increased every few days. Such treatment continued for a few weeks produces in the dog a diabetic state which persists indefinitely although the pituitary injections are discontinued. The only discoverable lesion is in the *islets*, which show pathological changes of varying degrees of severity ranging from depletion of the β cells to complete disappearance of the islets and their replacement by hyaline tissue. These animals differ from depancreatized dogs in several respects; in particular they can survive for long periods without insulin treatment. It is thought that the pituitary extracts first stimulate the islet cells (perhaps by producing hyperglycæmia) and then exhaust and finally destroy them, producing insulin lack. The experiments suggest the possibility that clinically, temporary pituitary disorder may produce a permanent islet lesion and consequently lasting diabetes.

(4) METATHYROID DIABETES.—Prolonged thyroid treatment likewise damages the islets (again perhaps because of the hyperglycæmia induced), but the injury is less severe than that induced by anterior pituitary extracts or by alloxan. Thyroid treatment will thus only induce a permanent diabetic state, *i.e.* one which outlasts its administration, if most of the pancreas is previously removed. The animal is then left without a reserve of islet tissue; if the remaining islet tissue is now injured by thyroid administration, a persistent insulin deficiency and diabetes result.

(5) ALLOXAN poisoning (p. 910).

(6) HIGH CARBOHYDRATE DIET.—It is claimed that ingestion of large amounts of carbohydrate over long periods may initially stimulate and finally exhaust and damage the islets.

Clinical Diabetes Mellitus.[2]—In the past, owing to the similarity between the metabolic changes in human diabetes mellitus and in the pancreatectomized animal it was generally supposed that the clinical condition was always due to a pancreatic lesion. Post-mortem examination of the pancreas gives only partial support to this view. Gross macroscopic changes in the pancreas are rare; œdema of the β cells of the islets has been demonstrated on microscopic examination a few hours after death, and regarded as evidence of deficient insulin secretion. Often, however, *no lesions can be found in the islets*. Soskin has said that *clinical* diabetes mellitus is a

[1] These results show that in man after pancreatectomy the insulin requirements are less than in some cases of clinical diabetes mellitus. It was pointed out on p. 910 that recently a blood glucose *raising* factor (called "hyperglycæmic factor") has been isolated from impure insulin and from the islets. These findings suggest the possibility that the hyperglycæmia in pancreatic disease may sometimes be due not only to insulin lack but also to excess of a blood glucose raising factor of pancreatic origin.

[2] Himsworth, *Lancet*, 1949, i, 465.

disease of unknown ætiology, thus emphasising our ignorance of the causal agent or of the site of the anatomical lesion in many cases; but whatever the site of the lesion in clinical cases the *functional* result is always to produce a *metabolic disturbance similar to that caused by insulin lack.* This view is confirmed by the finding that injection of insulin ameliorates the metabolic and clinical state in almost every case of clinical diabetes mellitus, whatever the cause may have been.

Clinically the diabetic syndrome may be due to:

(i) *Pancreatic islet deficiency:* unless there is evidence to the contrary, clinical cases of diabetes mellitus are always attributed (perhaps wrongly) to insulin lack. The clinician likes to restrict the term "diabetes mellitus" to this class of case.

(ii) *Hyperpituitarism: e.g.* in acromegaly (p. 941), or Cushing's syndrome (p. 965).

(iii) *Hyperthyroidism, i.e.* in Graves' disease (p. 991).

(iv) Increased secretion of *adrenal cortex* hormones (cortico-metabolic syndrome (p. 967)).

Metabolism in Diabetes Mellitus (Experimental and Clinical).—1. Changes in Carbohydrate Metabolism.— The changes described below can be readily deduced from the known actions of insulin.

(1) Glycogen Formation.—This is depressed in the liver and muscles:

(i) The resting liver glycogen level is low.

(ii) The resting skeletal muscle glycogen level is likewise diminished. In the normal animal the decrease in muscle glycogen content which occurs after tetanic stimulation is made good rapidly and completely; after pancreatectomy the restoration of the exhausted muscle glycogen is very much delayed.

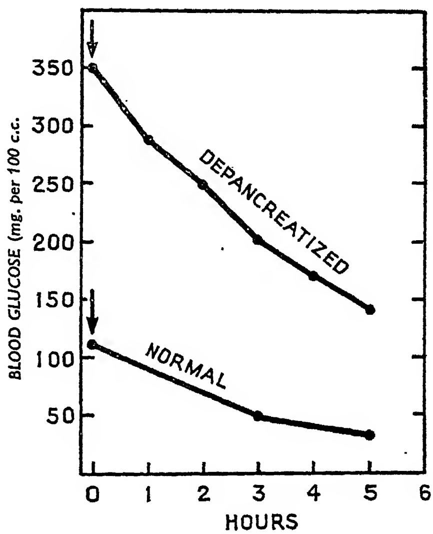


Fig. 585.—Effect of Removal of Liver on Blood Glucose in Normal and Depancreatized Animal. (Mann and Magath, *Arch. int. Med.*, 1923, *31*, 797.)

(2) Carbohydrate "Utilization" and Dissimilation by the Tissues.—There has been much discussion about this matter; at different times it has been argued that after pancreatectomy carbohydrate "utilization" or dissimilation is abolished, or is impaired, or is unaffected.

(i) It is certain that carbohydrate "utilization" is *not abolished.* This is proved by the studies on the eviscerate depancreatized preparation described on p. 911. Again, if the liver is removed in a normal animal the blood glucose falls; the rate of fall is a rough measure of the rate of carbohydrate "utilization" by the tissues. If the liver is removed in a *depancreatized* animal, the blood glucose (which is initially abnormally high) likewise

falls rapidly (Fig. 585), proving that glucose "utilization" is continuing on a considerable scale.

(ii) The rate of glucose dissimilation to CO_2 has been directly determined in diabetic animals using tracer methods. In the dog (which in its metabolic behaviour closely resembles man), the rate of glucose dissimilation as already noted is reduced by pancreatectomy to 40% of normal. Administration of insulin restores the rate of dissimilation to normal.

(3) Transformation of Glucose to Fatty Acids and Fat.—The conversion of glucose to long chain fatty acids in the liver and of glucose to fat in adipose tissue almost ceases. It had been supposed at one time that in the diabetic, fat is converted on a considerable scale into blood glucose; this is not the case. What happens is that dietary glucose which would normally have been removed and stored as fat now remains in the body fluids and thus raises the blood glucose level.[1]

(4) Hyperglycæmia. Rôle of Neoglucogenesis.—The blood glucose level is abnormally high, *e.g.* 200 mg-% or over, even in the fasting state. The hyperglycæmia is due to (i) decreased withdrawal of glucose from the body fluids (see (1)–(3) above) and (ii) to ill-regulated activity of the liver with respect to neoglucogenesis. The fundamental importance of the liver is shown by the fact that removal of the liver in the depancreatized animal produces fatal hypoglycæmia. In the fasting *normal* animal the liver manufactures new glucose from amino-acids (formed by the hydrolysis of protein) and secretes it into the blood in amounts sufficient to maintain the normal blood glucose level. In the diabetic state on the other hand the liver manufactures and secretes unnecessary "new" glucose into the blood and thus aggravates the hyperglycæmia. It is uncertain whether the *absolute* amount of glucose formed by the liver is greater than in the normal animal; but whatever the amount it is obviously more than is needed. As much of the blood glucose is excreted in the urine there is a considerable loss of energy. In the fasting diabetic the "new" glucose, which is wastefully made by the liver comes from tissue protein reserves; marked loss of body weight therefore occurs.

(5) Respiratory Quotient.—The respiratory quotient is lowered in the diabetic from the normal value of 0·8–0·85 (on a mixed diet) to values in the region of 0·7. This has been interpreted as signifying that fat alone is being dissimilated and not carbohydrate, *i.e.* as giving support to the extreme "non-utilization of carbohydrate" view which has been proved to be incorrect. But as explained on p. 373, the interpretation of the respiratory quotient is beset with difficulties; as a rule, far from the R.Q. increasing our knowledge of metabolic processes, a detailed study of the metabolic processes usually has to be undertaken to account for the R.Q. observed. The low R.Q. is probably due to several factors:

(i) The decreased dissimilation of glucose (R.Q. of glucose=1·0).

(ii) The increased dissimilation of fats (R.Q. of fat=0·7).

(iii) More fat is converted in the liver into ketones than the tissues can use; the excess ketones are excreted in the urine. The conversion of say, 1 mol. of palmitic acid into the equivalent 4 mols. of acetoacetic acid (if this is *not* dissimilated further) involves the utilization of 7 mols. of O_2 without any corresponding evolution of CO_2. Thus:

[1] Stetten, *Recent Progress Hormone Research*, 1949, *4*, 189.

$$\underset{\text{palmitic acid.}}{C_{16}H_{32}O_2} \xrightarrow{+7O_2} \underset{\text{acetoacetic acid.}}{4CH_3COCH_2COOH}$$

The measured R.Q. is the ratio of $\frac{CO_2 \text{ evolved}}{O_2 \text{ consumed}}$ by the *entire metabolism of the body*. The conversion shown in the above equation increases the denominator only and so lowers the R.Q.

(6) Glycosuria.—Owing to the hyperglycæmia the tubule reabsorption

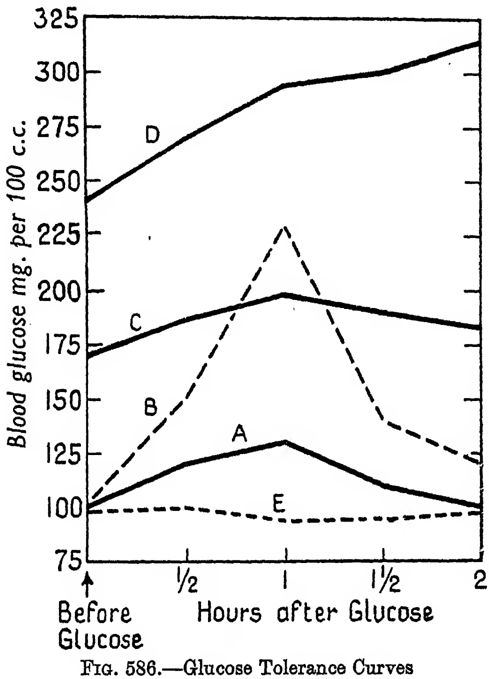


Fig. 586.—Glucose Tolerance Curves in Various Conditions.

(Drawn from data in Beaumont and Dodds, *Recent Advances in Medicine*.)

n each case the blood glucose was determined before and after the ingestion of 50 g. of glucose. A=Normal subject B=Mild diabetes; normal blood glucose during fasting. C=Severe diabetes; blood glucose when fasting. D=More severe diabetes. E=Renal glycosuria.

maximum (Tm_G) for glucose is exceeded (p. 927); in other words, the glucose filtered out of the blood into Bowman's capsule is not completely reabsorbed (as it is normally) by the tubular epithelium back into the blood; glycosuria results. Owing to the extensive reabsorption of water which always takes place in the tubules, the concentration of sugar in the urine rises to levels far above that in the blood, even, for example, up to 5 g-%. The osmotic pressure exerted by the high concentration of glucose in the urine is considerable and hampers the reabsorption of water; so more water than normally is

or more. Subsequently a very slow fall sets in, so that many hours may elapse before the fasting blood glucose level is regained.

The initial hyperglycæmia and the shape of the post-ingestion curve are due to lack of insulin and the consequent failure of the body to respond to the increase in the body-fluid glucose. The final fall of the blood glucose is due mainly to the excretion of glucose in the urine.

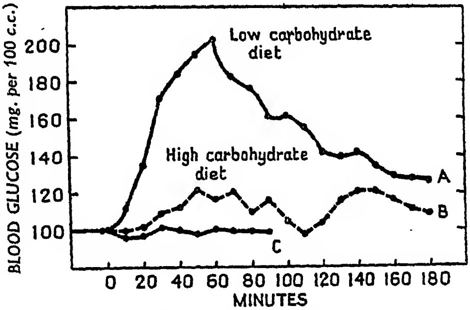


FIG. 587.—Effect of Carbohydrate Content of previous Diet on Glucose Tolerance Curve. (Himsworth, *Lancet*, 1939, ii.)

Diet A = 50 g. carbohydrate, 80 g. protein, 240 g. fat.
Diet B = 500 g. carbohydrate, 80 g. protein, 40 g. fat.
C = Control (cf. Fig. 588.)

2. **Changes in Fat Metabolism.**—(1) NON-TRANSFORMATION OF GLUCOSE INTO FATTY ACIDS AND FAT.—See p. 920.

(2) EXCESSIVE KETONE FORMATION IN THE LIVER.—Ketones are formed from fat in the liver and turned out into the blood at a faster rate than the tissues can use them; the blood ketone level is thereby raised (ketosis, ketæmia) and ketone bodies are excreted in the urine (p. 924). The excessive ketogenesis is related to the fact that the liver is depleted of glycogen and

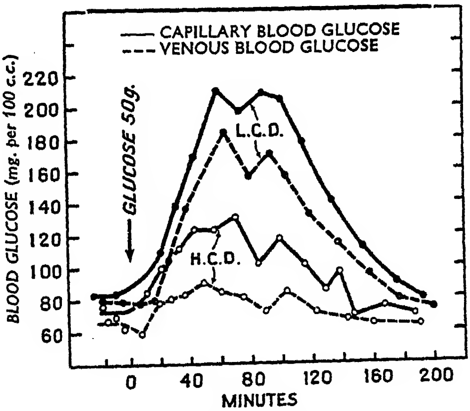


FIG. 588.—Effect of Carbohydrate Content of previous Diet on Peripheral Uptake of Glucose by Tissues. (Himsworth, *Lancet*, 1939, ii.)

Note the areas enclosed by the capillary-venous curves are similar in both cases.
L.C.D. = previous low carbohydrate diet. H.C.D. = previous high carbohydrate diet.

laden with fat. Even in the fasting state the ketosis persists, the ketone bodies being formed from fat which is mobilized in large amounts from the fat reserves.

The rôle of the liver in blood ketone regulation is as important as its rôle in blood glucose regulation. After hepatectomy the blood ketone level decreases as rapidly in the depancreatized as in the normal animal. If the liver is damaged in the diabetic patient both the ketosis and the hyperglycæmia are reduced. Though the diabetic manifestation are less marked the patient is worse because he now also suffers from the serious condition of hepatic failure (p. 869).

(3) KETONE UTILIZATION.—The power of the tissues to utilize ketones is unimpaired in diabetes. As less glucose is used a larger proportion of the energy needs of the body is supplied by ketones (or complete oxidation).

(4) The *blood fat* (in the form of triglyceride) rises (lipæmia). The lipæmia indicates that fat is " called out " of the depots *into* the blood faster than the liver can withdraw it *from* the blood; the development of fatty liver indicates that the liver is taking up fat from the blood faster than it can dispose of it by conversion into ketones.

(5) The *blood cholesterol* level also rises, even up to 1·5% in extreme cases.

(6) The *glycerol* ($CH_2OH.CHOH.CH_2OH$) of neutral fat is converted normally into glucose (in the liver).

RESULTS OF KETOSIS.—The ketone bodies are acetoacetic acid (the substance primarily formed) and β-hydroxy-butyric acid; these acids are interconvertible. Acetoacetic acid breaks up in the blood and the urinary bladder to form acetone which is eliminated in the expired air (giving it its characteristic smell) and in the urine (see p. 869).

The presence of abnormal amounts of acetoacetic and β-hydroxy-butyric acid in the blood tends to induce a " metabolic " acidæmia (p. 99). The usual compensatory reactions occur (p. 102):

(i) The acids are initially neutralized by the $NaHCO_3$ of the plasma.

$$H.Ak + NaHCO_3 \longrightarrow Na.Ak + H_2O + CO_2$$

(Ak is the acid ketone body under consideration.)

(ii) There is compensatory hyperpnœa which eliminates (*a*) the extra CO_2 liberated by the above buffering reaction and (*b*) additional CO_2 as well so as to get rid of enough extra CO_2 to make up for the decrease in the plasma alkali ($NaHCO_3$). The ratio determining blood reaction $\frac{(NaHCO_3)}{(H_2CO_3)}$ is brought back approximately to normal; the numerator (bicarbonate) is primarily reduced by the buffering reaction; the denominator (H_2CO_3) is secondarily reduced by the compensatory overventilation (p. 396).

(iii) There is excretion of a highly acid urine (p. 94).

(iv) There is increased NH_4^+ excretion: more NH_3 is made by the renal cells to neutralize the abnormal acids (p. 97).

Direct measurement of the H^+ ion concentration of the blood reveals little change in most cases of diabetes, proving how effectively compensation of the acidæmia is carried out.

3. **Changes in Protein Metabolism.**—Normally the non-nitrogenous residues of the amino-acids are converted into glucose (mainly) and ketones (to a small extent) in the liver. In the diabetic, neoglucogenesis from protein takes place excessively, and this overproduction largely contributes to the

hyperglycæmia; there is also some excess ketogenesis from protein which aggravates the ketosis (cf. pp. 887 *et seq.*)

In the fasting diabetic the tissue proteins are hydrolyzed rapidly to produce large amounts of amino-acids which are converted into glucose and ketones, much of which is excreted in the urine.

Effect of Diet in Diabetes.—Extra dietary carbohydrate increases the hyperglycæmia and the glycosuria (as occurs also when glucose is given); dietary protein is mainly converted into glucose and partly into ketones; dietary fat increases the ketosis only.

4. **Diabetic Coma.**—The onset of coma is foreshadowed by headache, nausea, vomiting, obstinate constipation, and abdominal pain. The breathing becomes deep and sighing ("air hunger"), and consciousness is gradually lost. If the alveolar CO_2 falls below 2%, coma may supervene within 24 hours. The factors producing the coma are: (i) Accumulation of the poisonous acetoacetic acid. (ii) The acidæmia which may be present. (iii) Failure of the kidneys: albuminuria, casts, diminished excretion of urine, and high blood urea may be found. There may be complicated changes in water and ionic balance. (iv) Circulatory failure is present, and the blood pressure is very low; this may be partly the result of a decrease in the plasma volume.

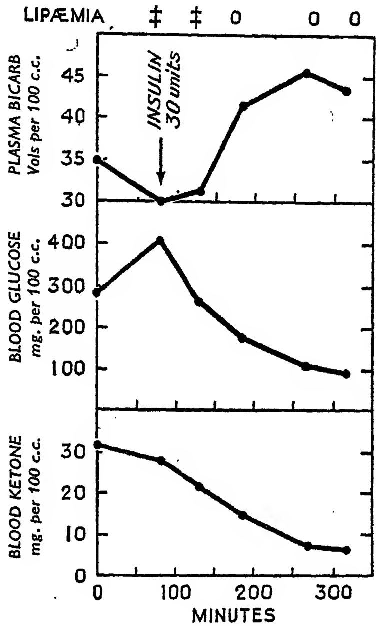


Fig. 589.—Action of Insulin in Case of Severe Diabetic Patient. (Davies, Lambie, Lyons, Meakins and Robson. *Brit. med. J.*, 1923.)

Ordinates from above downwards are: Plasma bicarbonate in vols/100 cc. Blood glucose in mg/100 cc. Blood ketones in mg/100 cc.
At arrow inject 30 units of insulin.
Note: Rise of plasma bicarbonate from 30 to 45 vols/100 cc. Fall of blood glucose from 400 to 100 mg/100 cc. Fall of blood ketones from 32to 8 mg/100 cc. Disappearance of lipæmia.

Action of Insulin in Diabetes Mellitus.—The administration of insulin to the diabetic restores him to normal:

(i) The inflow and outflow of glucose from the body fluids become normal. The rate of dissimilation of glucose, its conversion into glycogen, fatty acids, and fat, and the rate of neoglucogenesis by the liver become normal. The blood sugar falls to normal.

(ii) Ketone formation in the liver and the excessive mobilization of fat cease; as the ketosis disappears the reactions to acidæmia are dispensed with.

(iii) The R.Q. returns to normal.

(iv) The urine becomes normal (the glycosuria and ketonuria disappear) (Fig. 589).

Administration of Insulin.—Insulin can be administered clinically (usually by subcutaneous injection) in several different forms:

(1) *Regular Insulin :* an aqueous solution of crystalline or other purified insulin. It is rapidly absorbed from the subcutaneous tissues and lowers the blood glucose after a short latent period ; the effect passes off in a few hours (Fig. 590). A substantial proportion of the injected insulin is probably excreted.

(2) *Zinc-Insulin :* the addition of appropriate amounts of zinc (as a soluble salt, *e.g.* $ZnCl_2$) delays the absorption of insulin from the skin by a mechanism which is unknown ; its action is consequently prolonged.

(3) *Protamine-Insulin :* a clear solution of a protamine (a basic protein extracted from nuclei and containing much arginine and lysine) is added to an insulin solution ; the protamine unites chemically with insulin (as it does with other proteins). The protamine is dissolved in a suitable buffer solution of such composition that the pH of the mixed insulin and protamine solutions is 7·2. At this reaction the protamine-insulin compound is insoluble and forms a flocculent precipitate. The suspension is injected ; insulin is

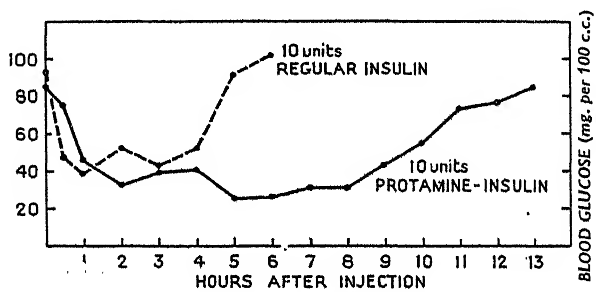


FIG. 590.—Effect of Regular Insulin and of Protamine-Insulin on Blood Glucose. Note the much more prolonged effect of protamine-insulin. (Modified from Best and Kerr.)

slowly liberated, slowly absorbed, and produces a prolonged effect (cf. androgens, p. 1111). The advantages of this method are : (i) The duration of the antidiabetic effect is, as stated, longer (Fig. 590). The number of injections required is therefore reduced. (ii) The blood glucose curve during the 24 hours is more regular and does not show a high peak after a meal or a marked depression after an injection. There are consequently fewer unpleasant hypoglycæmic reactions. (iii) The ketosis is better controlled.

(4) *Zinc-Protamine-Insulin.*—If a zinc chloride solution is added to a protamine-insulin suspension the latter becomes stabilized and remains in a satisfactory state in vials without sedimenting out, for months. The presence of zinc also facilitates the formation of the protamine-insulin compound.

(5) *Regular Insulin together with a " Retard " Insulin* (*e.g.* (2), (3) or (4)).—This combines the advantages of the immediate response to the regular insulin solution with the prolonged action of the more slowly absorbed forms. The two preparations must be injected from *separate* syringes at separate sites.

Abnormal Response to Insulin.—(1) INSULIN HYPERSENSITIVITY.—This condition is present when a severe prolonged hypoglycæmia (which may be fatal) results from a dose of insulin which produces a trivial blood glucose fall in the normal animal. It occurs experimentally after removal of the

glands which antagonize the action of insulin, *i.e.* after hypophysectomy (p. 938) or adrenalectomy.

Insulin hypersensitivity is present in clinical cases of hypopituitarism (*e.g.* Simmonds' disease) and of adrenal cortex deficiency (Addison's disease).

(2) INSULIN RESISTANCE.—Usually the subject under observation is a diabetic whose hyperglycæmia was previously kept in check by small doses of insulin ; with the development of insulin resistance, doses of, for example 200 units or more may fail to lower the blood glucose. The possible causes are :

(i) *Poor absorption* of insulin from the subcutaneous tissues (into which insulin is usually injected) ; such cases respond normally to insulin when it is administered intravenously.

(ii) Excessively *rapid destruction* of insulin : there is no evidence that this occurs.

(iii) *Overactivity of the insulin antagonists, i.e.* anterior pituitary, adrenal cortex, thyroid.[1]

(iv) *Abnormal condition of the liver* (due to "fatty state," action of toxins, or fever) so that it does not respond normally to its hormonal regulators, *i.e.* hepatic glucogenesis is not "screwed off" as effectively as usual by insulin.[2]

(v) In some cases there is evidence of the presence of an *anti-insulin* substance in the patient's blood ; injection of the "resistant" patient's serum plus insulin then produces a smaller fall of blood glucose in a test animal than an injection of normal serum plus insulin.

Glycosuria.[3]—The term glycosuria refers to the presence of sugar in the urine in sufficient amounts to reduce alkaline solutions of copper salt.[4]

The normal arterial blood glucose level is 80–180 mg.-%, according to circumstances. Glucose is filtered out from the plasma into Bowman's capsule ; the glucose concentration in the tubular fluid, therefore, is initially identical with that in the plasma. The normal average glomerular filtrate volume is 120 c.c. per minute. The amount of glucose filtered per minute must vary with : (i) plasma glucose concentration and (ii) the rate of glomerular filtration. The *maximal* amount of glucose that could be filtered per minute at the maximal normal blood glucose level is thus $\frac{120 \times 180}{100}$ *i.e.* 216 mg.

Glucose is transferred *actively* by the tubular epithelium from the lumen of the tubule back into the blood (reabsorption). There is an *upper limit* to the absolute amount of glucose which can be reabsorbed by the tubules in unit time ; this is called the *tubular absorption maximum for glucose* or the Tm_G ; its normal average value per minute is 300 mg. for women and 350 mg.

[1] The presence of the pancreatic "hyperglycæmic factor" (p. 918, footnote 1) may play a part in some cases.

[2] Those investigators who are never at a loss for a phrase to "explain" a phenomenon by describing it in other, more impressive words, have termed this non-reaction of the liver "hepatic refractoriness as regards the inhibitory response." There would seem to be some neglect of "Plain Words" here. (See Sir Ernest Gowers' *Plain Words*, Stationery Office.)

[3] Corcoran, *Cleveland Clin. Quarterly*, 1948, *15*, 186.

[4] Chromatographic examination of the urine shows that traces of glucose are normally present but the concentration is too low to give a reducing reaction. The presence of these normal traces is ignored in the discussion below.

for men. The blood glucose levels at which 300 and 350 mg. of glucose would be filtered out per minute from normal glomeruli are about 250 and 300 mg–% respectively.

This calculation does *not* give the minimum (*threshold*) blood glucose level at which glycosuria occurs. The reason is that some of the renal tubules have a *lower* maximal absorptive power than the average. When the amount of glucose filtered out in the glomeruli exceeds the Tm_G of these "weaker" tubules, detectable glycosuria occurs (although the other tubules have not yet exerted their full reabsorptive power).[1] Clinically, detectable glycosuria is commonly present when the blood glucose exceeds 180 mg–%, but individual variations are considerable. When the blood glucose level is such that the Tm_G of *all* the renal tubules is exceeded gross glycosuria occurs (Fig. 591). Glycosuria may thus result from:

(i) Increased filtration of glucose due to: (*a*) raised plasma glucose concentration; this is much the commonest cause; (*b*) increased volume of glomerular filtrate; this is very rare.

(ii) Decreased Tm_G; this may be due to a harmless individual variation, or result from the action of poisons or disease.

(1) Relation of Hyperglycæmia to Glycosuria.—Assuming a glomerular filtrate volume of 120 c.c. per minute and a Tm_G of 300 mg. per minute, *considerable* glycosuria will occur when the blood glucose exceeds 250 mg–%. If the Tm_G is greater or less than the average, the blood glucose level at which glycosuria is produced is correspondingly higher or lower (but see comment above).

Some patients with diabetes mellitus with a blood glucose level which is considerably raised do not have the expected glycosuria. Such a finding may be due: (i) occasionally to increased Tm_G; it is suggested that prolonged hyperglycæmia may stimulate the reabsorptive power of the tubular epithelium, *e.g.* to a maximum of 400–450 mg. per minute; (ii) much more commonly to *decreased glomerular filtration* resulting from renal disease or circulatory failure. In the latter case the absence of glycosuria is of serious omen.

(2) Relation of Variations in Tm_G to Glycosuria.—(i) *Renal glycosuria.*—This term is applied to a condition in which the Tm_G is considerably depressed without any other abnormality of renal function. Should the Tm_G fall to 120 mg. per minute, calculation shows that glycosuria would occur whenever the blood glucose level exceeds 100 mg–%. In renal glycosuria the glucose tolerance curve is flat (Fig. 586, curve E) owing to the leak of glucose into the urine.

(ii) *Phloridzin glycosuria.*—Phloridzin in large doses paralyses the power of the tubules to reabsorb glucose; *i.e.* the Tm_G may fall almost to zero. Glycosuria then occurs at all blood glucose levels, even during hypoglycæmia.

(iii) *Renal disease.*—In any condition in which the tubular epithelium is damaged a decrease in Tm_G occurs. Glycosuria may be a minor manifestation in cases of renal disease in which the lesion is tubular rather than glomerular.

(3) Alimentary Glycosuria.—If large quantities of polysaccharide (starch) or glucose are ingested, glycosuria does not occur except in people in

[1] The Table on p. 41 shows that before the Tm_G for the whole renal mass is reached small amounts of glucose may be passed in the urine.

whom the Tm_G is low; such people may excrete small amounts of glucose in the urine.

When large amounts of glucose are given by the mouth in man, the following results are obtained:

(i) With amounts of 150–200 g. of glucose, no glucose is found in a 24-hour specimen of urine.

(ii) When 300–500 g. are ingested, the *stomach* empties itself very slowly. Water is attracted osmotically into the stomach, gastric movements are inhibited, and the organ remains distended for hours with a large volume of glucose solutions. The stomach gradually empties itself into the small intestine where the glucose is absorbed in the usual way into the blood without an excessive rise of blood glucose and without the Tm_G being exceeded.

(iii) With amounts over 500 g. the limits of ingestion are reached as nausea develops and the glucose is vomited up.

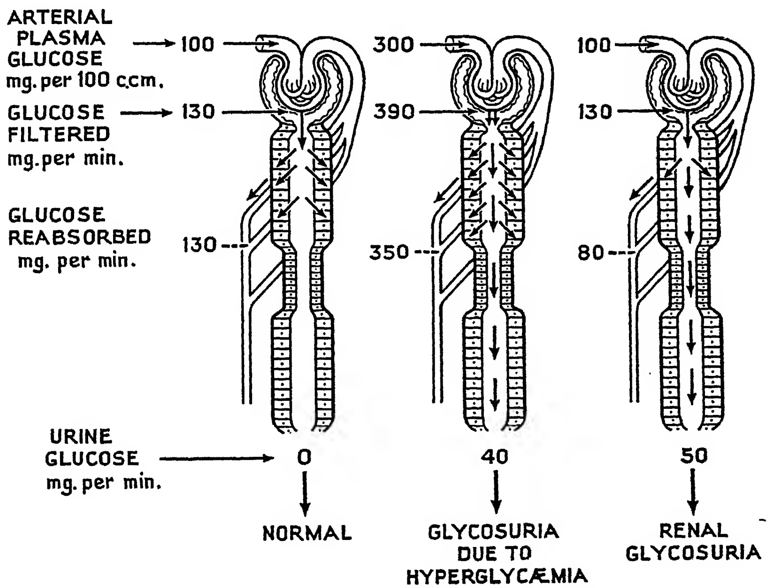


FIG. 591.—Mode of Production of Different Forms of Glycosuria. (After Corcoran, *Cleveland Clin. Quarterly*, 1948, *15*.)

PHYSIOLOGY OF ANTERIOR PITUITARY. CONTROL OF GROWTH AND METABOLISM. CLINICAL SYNDROMES

It is convenient to indicate here where the different aspects of the physiology of the pituitary are considered in this book.

1. POSTERIOR LOBE.—(i) Structure and functions with special reference to antidiuretic hormone, p. 41.

(ii) Rôle of oxytocin in parturition, p. 1091.

(iii) Relation to secretion of milk, p. 1094.

2. **Anterior Lobe.**—(i) Structure, p. 42.

(ii) The functions of the anterior pituitary are numerous. It controls growth, especially of the bones, muscles and viscera. It regulates the metabolism of carbohydrate, protein, and fat. It controls the growth, development, structural integrity and activity of the adrenal cortex, thyroid, testis, ovary and breast. Six hormones have been isolated from the anterior pituitary in a pure or almost pure state. They are: (*a*) growth hormone; (*b*) adrenocorticotrophic hormone (adrenocorticotrophin, ACTH); (*c*) thyrotrophic hormone (thyrotrophin, thyroid-stimulating hormone, TSH); (*d*) follicle-stimulating hormone (FSH); (*e*) luteinizing hormone (LH), identical chemically with interstitial-cell-stimulating hormone (ICSH); (*f*) prolactin (lactogenic hormone, lactogen, luteotrophin). All these hormones are simple proteins except (*d*) and (*e*) which are glycoproteins (containing a carbohydrate moiety).

(iii) Extirpation of pituitary, *infra*.

(iv) Control of the secretion of its hormones, p. 931.

(v) Relation to growth, especially of the skeleton (growth hormone); gigantism; acromegaly, p. 934.

(vi) Relation to general metabolism ("diabetogenic factor"), p. 936.

(vii) Relation to adrenal cortex (adrenocorticotrophin), p. 946.

(viii) Relation to thyroid (thyrotrophin), p. 980.

(ix) Relation to testis (interstitial-cell-stimulating hormone and FSH), p. 1113.

(x) Relation to ovary (gonadotrophins in the female, *i.e.* follicle-stimulating hormone, luteinizing hormone, prolactin), p. 1083.

(xi) Relation to breast (prolactin and other hormones), p. 1094.

(xii) Clinical syndromes, p. 940.

Extirpation of the Pituitary.—The results of complete hypophysectomy are due in part to depressed activity of other endocrines and in part to loss of the direct pituitary influence on many organs and tissues.

(i) In young animals growth ceases (*dwarfism*) (p. 934). In adults no skeletal changes occur.

(ii) The thyroid and adrenal cortex atrophy (pp. 980, 946).

(iii) The gonads remain infantile if the operation is performed before puberty; hypophysectomy in the adult causes atrophy of the gonads. Lack of testicular and ovarian secretion produces the expected effects on the accessory organs of reproduction and the secondary sexual characters (pp. 1068, 1107).

(iv) There are profound changes in metabolism due to lack of growth hormone, thyroxine, adrenal corticoids, and possibly other less well-defined factors (p. 931).

(v) Renal function is disturbed. Denervation or injury to the posterior pituitary produces diabetes insipidus (p. 49). Complete hypophysectomy depresses renal function generally. There is a marked decrease in glomerular filtration rate, renal plasma flow and maximal tubular excretion of diodrast; the diuresis which follows water drinking is greatly delayed.

Simmonds' Disease (Panhypopituitarism).—This is the classical clinical syndrome of grave anterior pituitary insufficiency in adults. It commonly follows severe post-partum hæmorrhage which somehow leads to extensive

pituitary necrosis. Atrophic changes are found in the gonads, thyroid and adrenal cortex. There is hypogonadism, *e.g.* decreased spermatogenesis, decreased secretion of testosterone and depression of the secondary sexual characters, impotence; amenorrhœa and sterility; loss of axillary and pubic hair; atrophy and pallor of the skin. There is mild thyroid deficiency, *e.g.* the B.M.R. falls to −20 (p. 980). There is mild adrenocortical deficiency but generally no significant changes in electrolyte and water balance (p. 946). The usual decrease in the eosinophil count in response to intravenous infusion of adrenaline may be absent (p. 951). The urinary output of neutral 17-ketosteroids is decreased markedly (*e.g.* to 0–2 mg./24 hr. (p. 953). Following an injection of insulin, the blood glucose falls normally, but recovery occurs very slowly ("hypoglycæmia unresponsiveness"). In some cases wasting (*cachexia*) has been noted, but this finding is by no means constant.

Control of Anterior Pituitary Secretions.[1]—The activities of the anterior pituitary are doubtless appropriately controlled to safeguard the interests of the rest of the organism. The present position however is that we know more about the effects produced by the pituitary than about the factors influencing the pituitary itself. Nothing is known, for instance about the control of the secretion of the growth hormone. But there is evidence that the secretion of certain of the anterior pituitary hormones is regulated by nervous and humoral mechanisms.

(1) Nervous Control.—Direct or reflex stimulation of the hypothalamus causes secretion of adrenocorticotrophin (p. 948), and of gonadotrophins (p. 1084); possibly secretion of thyrotrophin also occurs. Yet very few nerve fibres have been traced into the anterior pituitary and even these probably only supply the blood vessels. We have, therefore, the paradoxical situation of the nervous control of a gland which has no significant nerve supply. The explanation may be that the anterior pituitary has a double blood supply: (i) the usual systemic supply; (ii) a special *portal circulation* from the hypothalamus. The arteries which supply the median eminence break up locally into capillaries which drain into veins; these travel along the stalk of the pituitary to reach the anterior lobe where they break up into a second capillary net. The hypothalamo-hypophyseal veins are thus *portal* veins which could carry "products of hypothalamic activity" to the anterior lobe. It is suggested that the hypothalamus releases one or more chemical transmitters which are carried away in the blood of these portal veins to the anterior lobe, so regulating its activity. If these portal veins are destroyed the nervous control of the anterior lobe is abolished. The nervous control of the anterior lobe is thus due to unidentified humoral agents which travel a short distance in the blood; these agents may be contrasted with the chemical transmitters of the impulse in efferent nerves which act locally and the hormones of the ductless glands which are distributed by the circulation *all* over the body. The transmitters acting on the anterior pituitary may possibly be both adrenergic and cholinergic. It is interesting to note that adrenaline stimulates the secretion of ACTH.

(2) Humoral Control.—In several instances it can be shown that the activity of the anterior pituitary is regulated by the blood level of hormones secreted by glands which are under pituitary control; this mechanism is

[1] Harris, *Brit. med. J.*, 1951, ii, 627.

involved in the control of the secretion of thyrotrophin, gonadotrophins and ACTH.

(i) Thyrotrophin stimulates the secretion of thyroxine by the thyroid; the level of blood thyroxine in turn regulates thyrotrophin secretion: an increase in blood thyroxine depresses the release of thyrotrophin, while a decrease of blood thyroxine stimulates the release of thyrotrophin. Both pituitary and thyroid form part of a mutually interacting circuit, the activity of both glands being adjusted to maintain the normal state of the body (p. 979).

(ii) The level of blood œstrogen regulates the secretion of FSH (p. 1085).

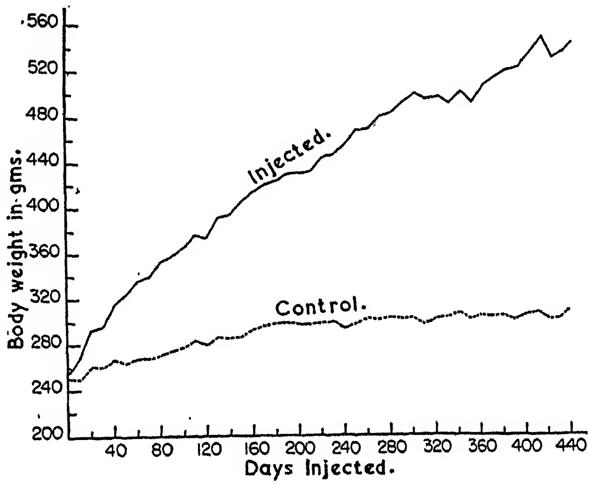


FIG. 592.—Effect of Growth Hormone on Growth of Adult Female Rats. (Li and Evans, *Recent Progress Hormone Research*, 1948, *3*, 5.)

Injected animals received *daily* 0·4 mg. of growth hormone for 23 days; 0·6 mg. for next 68 days; 1·0 mg. for next 33 days; 1·5 mg. for next 115 days; 2·0 mg. for next 193 days. No injections on Sundays. Controls received no injection.

(iii) The blood level of the adrenal corticoids may regulate the rate of secretion of ACTH (p. 948).

(iv) Testosterone has little effect on the anterior pituitary; the "backlash" (or "feed-back" as the physicists like to call it) from the testis to the pituitary may be due to an unidentified factor (p. 1114).

(v) Removal of the thyroid, adrenals, ovaries or testes produces histological changes in the anterior pituitary.

Relation of Anterior Pituitary to Growth and General Metabolism.—Three of the six hormones which have been isolated in pure or almost pure form, are related to growth and metabolism. They are: growth hormone (somatotrophin); adrenocorticotrophin (ACTH); thyrotrophin (thyroid-stimulating hormone, TSH).

(i) *Growth hormone:* its actions are fully described below.

(ii) *Adrenocorticotrophin:* this hormone acts by regulating the growth,

integrity and secretory activity of the adrenal cortex (p. 946). It is discussed below in so far as it is concerned with hypophyseal diabetes and diabetes mellitus.

(iii) *Thyrotrophin:* this hormone acts by regulating the growth, integrity and secretory activity of the thyroid (p. 980).

Fig. 593.—Hyperpituitarism with Giant Overgrowth. On the right is a normal subject for comparison (height 5 ft. 8 in.). On left is a case of hyperpituitarism with giant overgrowth. Note the narrow chest, large joints, hypotrichosis (defective growth of hair on body) and the large size of the hands. Harvey Cushing's case. (Vincent *Internal Secretion and Ductless Glands*, Arnold.)

(iv) Injection of certain anterior pituitary extracts produces a metabolic disturbance resembling diabetes mellitus (so-called *hypophyseal diabetes.*) The active agents are referred to collectively as the *diabetogenic factor.* A similar disturbance can be produced in some species by pure growth hormone and in others by ACTH. Possibly other substances which have not yet been identified with certainty may play a part also (*e.g.* glycotropic factor).

Relation of Anterior Pituitary to Growth. Growth Hormone.[1]— This hormone has been prepared in the pure form; it is a simple protein (mol. wt. 45,000). It is an indispensable component of the complex chemical systems present in the tissues which determine normal growth and differentiation.

1. **Experimental Evidence.**—(i) After hypophysectomy in young animals growth ceases; injection of growth hormone restores the rate of growth to normal.

(ii) If the hormone is injected into intact immature rats the rate of growth is accelerated for periods as long as 450 days.

FIG. 594.—Facial changes in Acromegaly.

A. Woman (age 29) before the onset of the disease. B. Same woman (aged 36), with fully developed disease. (After H. Zondek, *Krankheiten der Endokrinen Drüsen.*)

(iii) If injected into adult rats, overweight animals are produced (Fig. 592). The skeleton is larger and heavier, and the principal viscera are enlarged. The degree of growth obtained is related to sex: it is least marked in normal males, greater in normal females, and further increased in both sexes by removal of the gonads.

(iv) In young normal dogs receiving these extracts, changes in the skeleton resembling those found in human acromegaly (*infra*) have sometimes been produced.

2. **Clinical Evidence.**—Tumours of the anterior lobe may be associated with over-secretion—*hyperpituitarism*, or with under-secretion—*hypopituitarism*. When the latter condition develops in young people, stunted skeletal development (dwarfism) is a common feature. With hyperpituitarism, excessive skeletal growth occurs, the exact character of which depends on the time of onset of the disease.

(1) GIGANTISM.—This occurs when hyperpituitarism develops *before* the union of the epiphyses of the long bones. The bones continue to grow in a

[1] Li and Evans, *Recent Progress Hormone Research*, 1948, *3*, 1.

regular manner and become excessively long. Most giants who are seven or eight feet in height are examples of this disorder (Fig. 593).

(2) Acromegaly.—When hyperpituitarism sets in *after* the union of the epiphyses, acromegaly (literally, enlargement of the peripheral regions) results (Fig. 594). The bony changes produced in this disorder are as follows: There is overgrowth of certain parts of the skeleton, namely, the lower jaw, the upper jaw, and the malar bones (*i.e.* the lower half of the face); there is, consequently, separation of the teeth. The hands and feet are enlarged. There is bowing of the spine (kyphosis); the antero-posterior diameter of the chest is increased. The bent back, the big hands reaching down to the knees, the protruding lower jaw, remind one of the apt descriptive phrase—" a reversion to the gorilla type."

3. **Mode of Action of Growth Hormone.**—The mode of action of growth hormone can now be considered in greater detail.

(1) The increase in body growth produced by the hormone is due to a direct action on the tissues and is independent of any increased secretion by other endocrine glands (*e.g.* adrenal, thyroid, gonads).

(2) The muscles and viscera grow as rapidly as the skeleton and the body as a whole.

(3) Action on Bones.—After hypophysectomy a "calcium barrier" or " closing membrane " appears which cuts off normal activity at the epiphyseal line with the result that growth ceases. On injecting growth hormone the membrane disappears; the cartilage cells at the epiphyseal line begin to proliferate and differentiate once more into bone and bone growth is resumed.

Membrane bones (*e.g.* those of the cranium) are unaffected by hypophysectomy; bony tissue anywhere that has already been laid down likewise undergoes no change.

Injection of growth hormone (extracted from glands of cattle) into human pituitary dwarfs has so far proved completely ineffective (possibly because of species differences).

(4) General Metabolic Changes.—Growth involves the retention of all the materials necessary for the formation of fresh tissue, *e.g.* amino-acids for protein deposition, minerals, and water. As might be expected growth hormone affects many aspects of metabolism.

(i) Growth hormone decreases nitrogenous excretion in the urine (Fig. 595); the retained N is built up into tissue protein.

(ii) The blood amino-acid level falls, indicating that the amino-acids are being withdrawn from the blood more rapidly than they enter it.

(iii) Serum alkaline phosphatase is increased: this enzyme may be related to protein synthesis as well as to ossification (p. 1002).

(iv) Reserve fat is decreased, though *structural* fat (the lipides which are an integral part of the cell architecture) is increased in amount.

(v) Growth hormone causes increased retention of calcium and phosphate. The level of serum inorganic phosphate is increased (as it is in growing children and in acromegaly).

(vi) The thymus like the rest of the body is stimulated to increased growth (the thymus is often enlarged in acromegaly).

(vii) The yield of milk in lactating animals is increased (p. 1094).

(5) Diabetogenic Action of Growth Hormone.—In certain circumstances, growth hormone injected into normal animals produces a condition

resembling diabetes mellitus; there is a rise of blood glucose and development of glycosuria; reserve fat is decreased; there is ketonæmia and ketonuria. According to Young, growth hormone is diabetogenic in the adult dog and cat but not in the puppy or kitten or in the pregnant or lactating adult dog or cat. It is not consistently diabetogenic in adults of other species studied. The diabetic effect is more readily produced after partial pancreatectomy (owing to deficiency of the antagonistic hormone, insulin) (cf. p. 938).

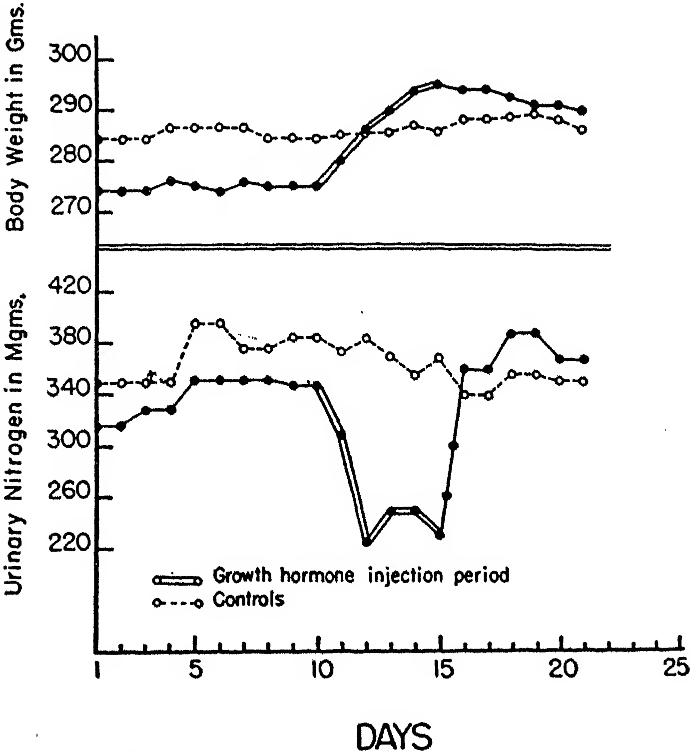


FIG. 595.—Effect of Growth Hormone on Body Weight and Nitrogenous Excretion in the Urine of Adult Female Rats. (Li and Evans, *Recent Progress Hormone Research*, 1948, *3*, 14.)

During the injection period the nitrogenous excretion in the urine was decreased (from about 340 mg. to about 220 mg. daily), and body weight increased (from 275 to 290 g.).

It is not at all easy (for the uninitiated) to understand how a hormone which promotes growth should also be able to produce diabetes, a condition which involves wasteful transformation of protein into glucose. The question is discussed again below.

Relation of Anterior Pituitary to Metabolism and especially to Diabetic Syndrome.[1]—This subject is complicated and confused. As already explained crude anterior pituitary extracts are diabetogenic. It

[1] Bennett and Evans, in *The Hormones*, N.Y., 1950, *2*, 405.

will be convenient for purposes of discussion to group all the diabetogenic substances in anterior pituitary extracts together and refer to them as though they constituted a physiological entity, the *diabetogenic factor*.

(1) ACTION OF "DIABETOGENIC FACTOR."—(i) In the *liver* it produces the opposite effects to those of insulin. Thus: (*a*) it increases the conversion of glucogens (glycogen and amino-acids) into glucose; (*b*) it increases the conversion of fatty acids into ketone bodies.

(ii) (*a*) In the *tissues* it inhibits the conversion of glucose into hexose-6-phosphate, thus preventing its introduction into the "metabolic machinery of the cell" (p. 914). This inhibitory action of anterior pituitary extracts is antagonized by insulin. (*b*) It stimulates the conversion of tissue protein into amino-acids. (*c*) It mobilizes depot fat.

(iii) It may directly *neutralize* insulin.

(iv) It may *damage the islets* (metahypophyseal diabetes, p. 918).

In some species many of these actions can be reproduced by growth hormone, in others by ACTH which acts by releasing adrenal corticoids.

(2) EFFECTS OF HYPOPHYSECTOMY IN NORMAL ANIMALS.—The main disturbances are due to loss of the diabetogenic factor (as defined above) which is supposed to be normally secreted steadily into the blood. Normally the action of the anterior pituitary factor and of insulin on the liver and tissues are nicely balanced, as shown below:

$$\text{Glucogens in Liver} \quad \underset{\text{Insulin}}{\overset{\text{Ant. Pituitary}}{\rightleftarrows}} \quad \text{Glucose in Blood}$$

$$\text{Glucose in Tissues} \quad \xrightarrow[\text{Anterior Pit. inhibits}]{\text{Insulin stimulates}} \quad \text{Hexose Phosphate}$$

In the absence of the pituitary factor, glucogenesis by the liver is depressed by the unopposed action of insulin and even if the diet is adequate, the blood glucose level falls by 20–30 mg-% below that in normal controls.

The effects of hypophysectomy on glucose uptake by the tissues have not been satisfactorily determined; it is probably increased, thus aggravating the tendency to hypoglycæmia.

In hypophysectomized animals, a period of *fasting* (which does not depress the blood glucose in intact animals) soon produces *hypoglycæmia* which develops rapidly and may be *fatal*. Protein given by *mouth* raises the low blood glucose level; *i.e.* the amino-acids derived from the food protein can be converted into blood glucose; food fat, however, is ineffective (Fig. 596). These results suggest that after hypophysectomy neither food fat nor depot fat can "spare" body glucose; *food protein, but not tissue protein, can be converted into glucose.* Now, during starvation in the intact animal, tissue protein is broken down on a large scale to give rise to amino-acids. As already mentioned the hypophysectomized animal can convert amino-acids, absorbed from the bowel, into glucose but it cannot maintain its blood glucose level when fasting. It seems, therefore, that after hypophysectomy tissue protein is not converted at the normal rate into amino-acids.

(3) Effects of Hypophysectomy in Pancreatectomized Animals.—Removal of the pituitary in pancreatectomized animals produces striking results, as was first demonstrated by Houssay. There is a marked improvement in the clinical state, and hyperglycæmia and glycosuria may practically disappear; ketosis is generally absent. Such a *Houssay animal*, as it is usually called, may live for months without special treatment, but its existence is always somewhat precarious. It becomes increasingly undernourished and finally dies, the longest recorded survival period being 9 months. The blood sugar in Houssay animals shows wide fluctuations. *Fasting* produces the same hypoglycæmic effect as in the hypophysectomized animal. After a *meal* the blood glucose rises to 200 or 300 mg-% and glycosuria may consequently develop. The degree of glycosuria depends on the species and the kind of food eaten; a carbohydrate- or protein-rich meal raises the blood glucose (Fig. 597); a high fat diet has no such effect.

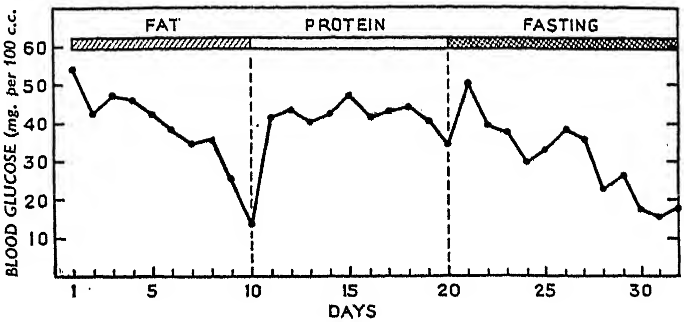


Fig. 596.—Action of High Fat Diet, High Protein Diet and Fasting on the Blood Glucose Level in the Hypophysectomized Dog. (After Soskin *et al.*, *Amer. J. Physiol.*, 1935, *114*, 110.)

High fat diet: 11 g. of olive oil per kg.
High protein diet: 11 g. of protein (as lean meat) per kg.
On the high protein diet the blood glucose is maintained at a level of 40–50 mg/100 cc.
On a high fat diet and during fasting much more severe hypoglycæmia develops.

The Houssay animal is very sensitive to: (i) the action of anterior pituitary extracts, which raise the blood glucose and produce glycosuria and ketosis; (ii) the action of insulin, which produces a striking hypoglycæmia which can be fatal.

The observations described above demonstrate that the diabetogenic factor is concerned with the *normal* regulation of metabolism. How its rate of secretion is adjusted in accordance with body needs is uncertain.

When the islets are removed the resulting diabetic syndrome is partly due to the unopposed action of the pituitary; if both the islets and the pituitary are removed, the two main regulators of carbohydrate, protein, and fat metabolism are absent, but on a suitably adjusted diet the animal remains fairly normal; but because it lacks the usual regulatory and compensatory devices, its condition is precarious and its blood glucose and general condition fluctuate widely with changes in the diet (Fig. 597).

(4) Effect of Adrenalectomy in Pancreatectomized Animals.—In some species adrenalectomy has the same effect as hypophysectomy in

ameliorating the diabetic symptoms. The adrenal corticoids thus have actions which are antagonistic to those of insulin (p. 945). As might be expected, injection of corticoids (e.g. cortisone) aggravates the symptoms of clinical diabetes mellitus.

(5) Glycotropic Factor (Young).—This substance (which has been extracted from the anterior pituitary) has *no direct action on the blood glucose*

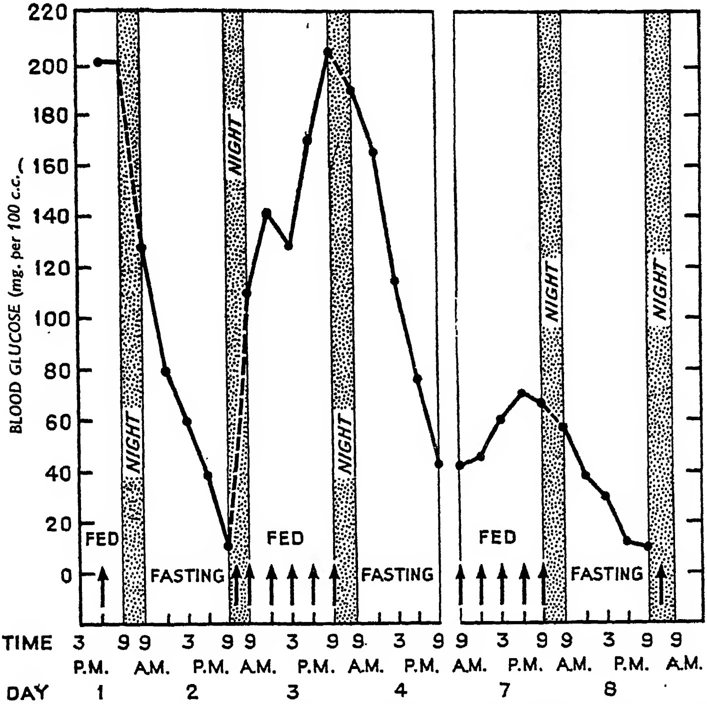


Fig. 597.—Influence of Size of Protein Intake and of Fasting on Blood Glucose Level of Dogs in which both the Pituitary and Pancreas had been removed. (After Soskin *et al.*, *Amer. J. Physiol.*, 1935, *114*, 110.)

The dog fasted on days 2, 4, and 8. It was fed on days 1, 3, and 7. On days 1 and 3 the protein intake was high (400 g. and 378 g.). On day 7 the protein intake was low (90 g.).

Note fall of blood sugar: (i) at night; (ii) by day when fasting; (iii) on a low protein diet.

level, but decreases the blood glucose lowering action of insulin (Fig. 598). Its absence is alleged to account in part for the enhanced insulin sensitivity of the hypophysectomized animal; the presence in excess of such a substance may account for some clinical cases of insulin resistance (p. 927).

(6) Relation of the Pituitary to Clinical Diabetes Mellitus.—From the above discussion it is clear that in considering clinical diabetes mellitus the anterior pituitary must be constantly kept to the fore as an important factor. It may *damage the islets* (p. 918); it may *diminish responsiveness to insulin* secreted in normal amounts by perfectly healthy islet

tissue (glycotropic factor) (p. 939); or it may actively *disturb carbohydrate, protein, and fat metabolism* in the presence of the normal pancreas (p. 937). In the pancreatectomized animal, or in a patient with primary islet deficiency, many of the symptoms are due to the unopposed action of the diabetogenic factor. The frequent clinical association of pituitary disease (*e.g.* acromegaly, Cushing's disease) with diabetes mellitus thus acquires an added significance.

Clinical Pituitary Syndromes.—Either increased pituitary secretion (*hyper*pituitarism) or diminished activity (*hypo*pituitarism) may occur.[1] A state of hypersecretion may later degenerate into a state of under-activity. Again, different functions of the gland may be affected simultaneously but in different directions. To such complex derangements of function the term *dys*pituitarism is applied.

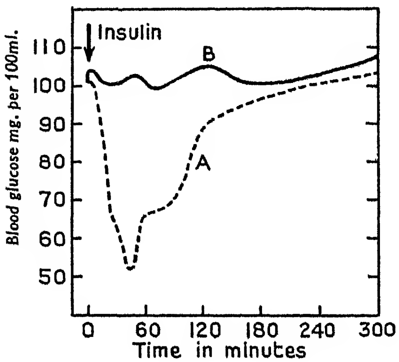


FIG. 598.—Influence of Glycotropic Factor of Anterior Pituitary on Action of Insulin on Blood Glucose. (Young, 1938.)

A represents normal animal; B, animal (rabbit) previously treated with anterior pituitary glycotropic factor. At the arrow inject test dose of insulin. Note that the factor abolishes the hypoglycæmic response to insulin.

(1) SIMMONDS' DISEASE.—See p. 930.

(2) MILDER SYNDROMES OF HYPOPITUITARISM.—The following changes are commonly found.

(i) Stunted skeletal development (*infra*).

(ii) Depression of sexual function.

(iii) Changes in skin and hair: in the male the beard and axillary hair are scanty, the pubic hair is of feminine distribution, the trunk and limbs are hairless. In children the skin is smooth and delicate; in the adult it is dry, wrinkled, and atrophic; the secretion of sweat is diminished.

(iv) Adiposity is fairly common, and there is a tendency for sudden fluctuations in weight to occur. In children the fat has a feminine distribution, but in adults the subcutaneous fat is grossly excessive in most regions.

(v) Mental backwardness.

(vi) Lowered basal metabolic rate.

(vii) Polyuria, which may be intense (due to lack of antidiuretic hormone, p. 49).

All or only certain of these features may be present in any individual case of this disorder.

When hypopituitarism develops in childhood it may assume the following forms:

(i) *Fröhlich's syndrome* (*dystrophia adiposo-genitalis*): several pituitary functions are disturbed; the children become stunted and stupid or quite idiotic. The genitals are hypoplastic, and there are diffuse deposits of fat.

(ii) *Lorain type:* mentally, these patients are normal; general metabolism is unaffected. But skeletal growth ceases and the secondary sexual characters

[1] Sheehan *et al.*, *Quart. J. med.*, 1949, *18*, 319; *Brit. med. J.*, 1950, i, 929.

do not appear. Even when they are of adult age they resemble attractive and graceful children.[1]

(3) DIABETES INSIPIDUS.—See p. 49.

(4) HYPERPITUITARISM.—In both gigantism and acromegaly an eosinophil adenoma of the anterior lobe of the pituitary is present (p. 44). The skeletal changes in these conditions were described on p. 934 and are due to over-secretion of the growth hormone. The other important features of *acromegaly* are as follows: (i) *Glycosuria* of extreme variability (some of these cases may die in diabetic coma); it is relieved by insulin. (ii) There is considerable *connective tissue hyperplasia*—the subcutaneous tissues of hands, feet, nose, scalp, and lips are markedly thickened; the skin is stiff and inelastic. Deep corrugations are present in the palms of the hands and on the scalp. (iii) *Sweating* is very profuse.

When the disease progresses, as it usually does, to hypo-activity of the gland, the clinical picture is altered. The skeletal changes are, of course, fixed and cannot be affected. Sexual activity diminishes—the males become impotent and, in the female, the menstrual flow ceases. There is excessive deposition of fat in the subcutaneous tissues. Gradually increasing weakness develops, and finally death results.

The tumour may press on adjacent structures, *e.g.* on the decussating fibres in the optic chiasma, causing loss of the temporal field of vision in each eye (Fig. 366, p. 580), or on one optic nerve, causing unilateral blindness.

(5) CUSHING'S SYNDROME.—See p. 965.

REGULATION OF GROWTH

Regulation of Growth.[2]—A distinction must first be drawn between growth and differentiation. *Growth* consists in an increase in the number and size of the cells of an organ or tissue; by *differentiation* is meant the transformation of a homogeneous group of cells into two or more groups of cells differing with respect to histological or physiological properties. The increase in the size of the fertilized ovum as a result of the initial stages of cell multiplication is a process of simple growth; the subsequent formation of the specialized layers of ectoderm, mesoderm, and endoderm is a process of differentiation. The processes of differentiation which lead to the appearance of the many specialized organs and tissues of the fœtus form one of

[1] Another classical syndrome hitherto attributed to juvenile hypopituitarism is the *Brissaud type*, exemplified by the "fat boy" (Joe) in Dickens' *Pickwick Papers*. It is more probable however that this condition results from a lesion of the hypothalamus. The outstanding symptoms in Joe's case were gluttony, gross obesity ("the leaden eyes which twinkled behind his mountainous cheeks leered horribly upon the food as he unpacked it from the basket") and frequent attacks of somnolence; the obesity was presumably secondary to the greed. "On the box sat a fat and red-faced boy in a state of somnolency. . . . Joe!—damn that boy, he's gone to sleep again.—Joe, let down the steps. . . . Mr. Winkle mounted to the box, the fat boy waddled to the same perch and fell fast asleep instantly." During the military exercise at Rochester "everybody was excited except the fat boy and he slept as soundly as if the roaring of cannon were his ordinary lullaby." "He's always asleep. Goes on errands fast asleep and snores as he waits at table." A similar condition has been produced by appropriate experimental hypothalamic lesions.

[2] White House Reports, *Growth and Development of Child*, Pts. I.–IV., New York, 1932.

the subject matters of embryology and will not be considered here. During the first three months of human intra-uterine life little actual growth takes place; this process is greatly accelerated during the later months of pregnancy. After birth the general growth curve shows four distinct phases (Fig. 599): (i) a rapid increase during infancy and especially in the first year (from 8 to 20 lb. in weight); (ii) a slow progressive growth from the age of 3 to 12; (iii) marked acceleration round about puberty, from the age of 12 to 15; (iv) a slow terminal increment. The growth curve shows distinct *sex* differences. Until just before puberty the boys are heavier; the earlier incidence of puberty in girls causes them to lead temporarily, but after the age of 15–16 the boys gain once more and maintain their advantage throughout adult life. These sex differences apply not only to growth as a whole, but also to that of the individual organs.

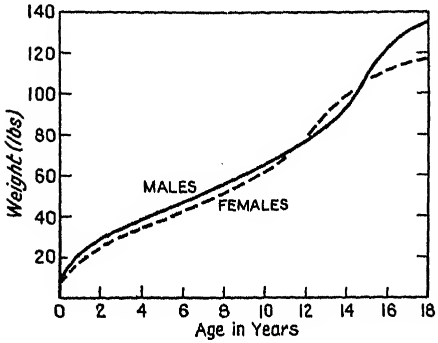


FIG. 599.—General Growth Curve in Males and Females. (*White House Reports*, 1932.)

The *general growth curve* applies to the skeleton as a whole, the muscles, and the viscera in the thorax and abdomen (Fig. 600, curve C). Certain parts of the body have distinctive growth curves. Three main specialized types are recognized:

(i) *Neural* type.—There is a rapid initial increase in size, so that the organs reach almost their final maximum in the first few years. This applies to the brain as a whole and its various parts, the spinal cord, the eye and other parts of the visual mechanism, parts of the auditory mechanism, and, naturally, to the size of the head (but not of the face) (Fig. 600, curve B).

(ii) *Lymphoid* type.—The lymphoid tissues including the *thymus* (and, be it noted, the *tonsils* and *adenoids*) normally reach their peak at puberty, after which they decline to the adult level (Fig. 600, curve A).

(iii) *Reproductive* type. — The gonads and the accessory organs of reproduction remain quiescent till puberty, when growth sets in and continues rapidly throughout adolescence (Fig. 600, curve D).

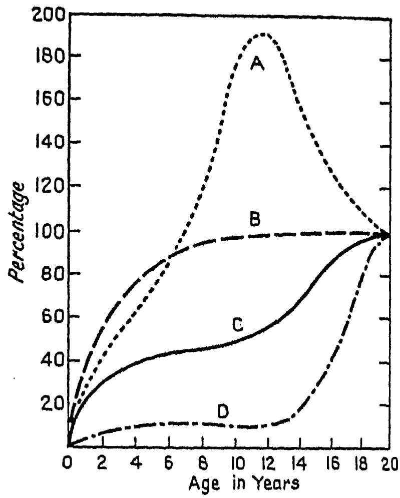


FIG. 600.—Growth Curves of different Types of Organs. (*White House Reports*, 1932.)

A. Lymphoid type. C. General type.
B. Neural type. D. Reproductive type.

Factors Controlling Growth.—The degree and rate of growth is influenced by many factors :

(1) The inherent properties of the fertilized ovum itself, which depend on the characters present in the chromosomes of the sperm and ovum. Growth thus varies in different races and in different members of a race. The internal and external environment of the body can at best bring out the growth potentialities to the full, but can normally do no more ; on the other hand, imperfections in the environment may lead to a failure to realize the maximum inherent growth possibilities.

(2) Ill-health from any cause temporarily depresses general growth, but during recovery the lost ground is made good again, more or less completely.

The factors influencing growth about which most is known are the *dietetic* and *endocrine.*

(3) *Rôle of Diet.*—On pp. 1051 *et seq.* is set out our present state of knowledge of the composition (qualitative and quantitative) of a satisfactory diet. If the diet is deficient in any way, growth will be impaired. It is highly likely that present so-called " normal " standards for height and weight do not represent the maximum potentialities, but are average findings in people who have not had a perfect diet. In many instances supplementary rations of milk and other high quality foodstuffs have produced higher growth standards in seemingly normal children

(4) *Rôle of Endocrines.*—The *growth hormone* of the anterior pituitary (p. 934) regulates the growth of the skeleton, muscles, and viscera. The three *gonadotrophic hormones* (p. 1083) regulate the growth (and differentiation) of the gonads; the *sex hormones* (pp. 1075, 1104) in their turn control the growth (and development) of the accessory organs of reproduction. The *thyroid* (p. 986) regulates general growth, presumably by its action on tissue metabolism ; pituitary *thyrotrophin* acts indirectly via the thyroid. The *parathyroid* glands (p. 1008) are related to the integrity of the skeleton rather than to its growth ; in certain tumours of the *adrenal cortex,* growth may be modified (p. 967). The special growth changes that occur during pregnancy in the uterus, vagina, breasts, and elsewhere, are controlled by hormones (for references see p. 1087).

THE ADRENAL CORTEX [1]

The adrenal gland is made up of two distinct organs, the *adrenal cortex* and the *adrenal medulla,* which differ in their histological structure, comparative anatomy, development, and functions.

Structure of Adrenal Gland.—(1) Adrenal Medulla.—The medulla consists of irregularly arranged cells, some polyhedral and others indefinite in shape, which contain granules of the internal secretion, probably adrenaline and nor-adrenaline. Oxidizing agents like dichromates (or iodates) convert the granules into a brown oxidation product which gives a characteristic colour to the cells ; for this reason the medullary cells are sometimes referred to as *chromaphil tissue.* The medulla has a rich sympathetic innervation. Its physiology is dealt with fully on pp. 723 *et seq.*

(2) Adrenal Cortex.—The *cortex* of the gland consists of columns of

[1] Noble, in *The Hormones,* N.Y., 1950, *2,* 65. Symposium, *Ann. N.Y. Acad. Sci.,* **1949,** ***50,*** **509–678.**

polyhedral cells with well-marked nuclei. The cells are arranged in three zones: (i) outer *zona glomerulosa* where the cells lie with their long axis parallel to the surface; (ii) middle *zona fasciculata* which consists of columns which lie vertical to the surface; (iii) inner *zona reticularis* where the cells are arranged irregularly, leaving wide blood spaces. The cortex has *no nerve supply*.

The cells of the cortex have a high lipide content; the chief constituent is *cholesterol*, mainly in the ester form, which is the precursor of the adrenal cortex hormones (*corticoids*). The lipide granules are readily stained with Sudan III (*sudanophil granules*). Sections stained in this way, obtained from

FIG. 601.—Histological Appearance of Adrenal Cortex Stained with Sudan III under Different Conditions. (Sayers and Sayers, *Recent Progress Hormone Research*, 1948, *2*, 81.)

Adrenal Cortex, 21-day-old male rats. Black and white reproduction of sections stained with Sudan III. Dark staining material is sudanophil substance, mainly (90%) cholesterol esters.

1. Control. Note deep staining of zona glomerulosa with narrow clear band deep to it. Zona fasciculata is deeply stained; zona reticularis is unstained.
2. Three hours after injection of 2 mg. of ACTH. Almost complete disappearance of sudanophil granules.
3. Twenty-four hours after a series of injections of ACTH over a period of three days. The cortex is hypertrophied and packed almost uniformly with sudanophil granules.

an animal under good environmental conditions, show a narrow clear area deep to the zona glomerulosa; the zona reticularis is also poorly stained (Fig. 601). The cortex is also rich in ascorbic acid which is possibly concerned with the synthesis of the corticoids. The size of the cortex, its histological appearance and chemical composition vary widely with different environmental conditions.

The physiology of the adrenal cortex is dealt with fully below.

COMPARATIVE ANATOMY.—In elasmobranch fishes no adrenal gland is present. Lying between the two kidneys is an elongated body, the *inter-renal body* which structurally closely resembles the mammalian adrenal cortex. Adjacent to the ganglia of the lateral sympathetic chain are masses of cells giving the chromaphil reaction, which together represent the adrenal medulla.

EMBRYOLOGY.—The adrenal cortex is mesoblastic in origin and develops from a thickening of the cœlomic epithelium on either side of the root of the mesentery. The adrenal medulla is developed from a mass of cells which first lie close to the root ganglia, and later migrate out to differentiate into the sympathetic nervous system and the chromaphil tissue.[1]

Effects of Adrenalectomy.—The adrenal cortex is essential to life. If both adrenal glands are completely removed death results in most species in 4 or 5 days unless accessory cortical tissue is present. If the medulla alone is destroyed or if a small fragment of cortical tissue is left in the body no serious effects are observed. Complicated disturbances follow on adrenalectomy.

(i) The animal loses weight rapidly and becomes weak and uncertain on its legs.

(ii) There are marked changes in fluid and ionic balance, which are in the main secondary to disturbed renal function.

(iii) The metabolism of carbohydrate, protein, and fat is deranged.

(iv) Death finally occurs chiefly from circulatory and renal failure.

Injection of adrenal cortex extracts into an adrenalectomized animal produces marked benefit or full recovery. The active agents are termed *corticoids*. All those that have been isolated in a chemically pure state are *steroids*. In addition a highly active amorphous material of unknown chemical composition (*amorphous fraction*) has been extracted from the gland. No one knows for certain which of the adrenal steroids is the actual internal secretion of the gland. Opinions differ as to whether the cortex secretes one, two, three or even more separate hormones. There is also a common tendency (which must be avoided) to regard the adrenal steroid which happens to be available in largest amount for experimental or clinical study as the natural hormone of the gland.

The chemistry and actions of the adrenal steroids are discussed on p. 957.

Functions of Adrenal Cortex.—The chief known functions are summarized below:

(1) (i) It acts on the renal tubular epithelium stimulating the reabsorption of Na^+ (and less regularly of Cl') into the blood thus conserving the Na^+ (and Cl') content of the extracellular fluids. It depresses the reabsorption of K^+. Disorders of ionic and fluid balance occur in states of decreased or increased cortical activity.

(ii) It acts similarly on the sweat glands and possibly on other glands (*e.g.* the salivary) decreasing the Na^+ (and more irregularly the Cl') content of the secretion.

(2) The adrenal cortex regulates *general* metabolism, *i.e.* that of carbohydrate, protein, and fat.

(i) Adrenal cortex extracts are *diabetogenic*.[2] Like the anterior pituitary

[1] ACCESSORY ADRENAL TISSUE.—Accessory adrenal tissue is found in various regions in mammals. Common situations for accessory *cortical* tissue are the retroperitoneum, the surface of the kidney and liver, the space between the testis and the epididymis in the male, and the vicinity of the broad ligament in the female. Accessory *chromaphil* tissue may be found in the sympathetic ganglia and in front of the abdominal aorta.

[2] As explained on p. 936, in some species the diabetogenic action is mainly due to growth hormone, in others it is mainly due to ACTH stimulating the secretion of adrenal corticoids.

diabetogenic factor they stimulate neoglucogenesis in the liver from amino-acids; they also inhibit the uptake of glucose by the tissues by preventing its conversion into hexose phosphate (p. 913).

(ii) Adrenal cortex extracts promote the breakdown of tissue protein to amino-acids.

(iii) They mobilize fat from the depots and increase its breakdown in the liver into ketone bodies.

(iv) Unlike anterior pituitary extracts they increase the deposition of glycogen in the liver.

(v) In the main the effects on metabolism of adrenalectomy resemble those of hypophysectomy. Over-secretion of the adrenal cortex produces a condition resembling mild diabetes mellitus.

(3) The adrenal cortex through its hormone enables the body to cope more effectively with adverse environmental conditions designated collectively states of *stress*. The exact way in which this beneficial influence is produced is not known (p. 947).

(4) The adrenal cortex controls the activity of the *lymphoid* tissues and the blood *eosinophil* level (p. 951).

Control of Activity of Adrenal Cortex.—The main regulator of the activity of the adrenal cortex is the anterior pituitary through its hormone *adrenocorticotrophin* (*ACTH*, *corticotrophin*). It controls the growth of the cortex, maintains the structural and functional integrity of its cells and stimulates its secretory activity. The methods of study include:

(i) Determination of the size of the adrenal cortex.

(ii) Histological examination for changes in the size of the individual cells and the distribution and amount of sudanophil granules (which are directly related to the cholesterol ester content).

(iii) Chemical determination of the concentration in the cortex of cholesterol esters (the precursors of the corticoids) and ascorbic acid (which may be involved in corticoid synthesis).

(iv) Changes in the organs, tissues and body fluids which are attributable to the action of secreted adrenal corticoids. Useful *clinical* indices are: the blood eosinophil count (p. 951); the Na^+ content of sweat (p. 952); the amount of neutral 17–ketosteroids and corticoids (so-called 11–oxysteroids) excreted in the urine (p. 952).

Action and Rôle of Adrenocorticotrophic Hormone (ACTH).—The action of ACTH in man is mainly discussed elsewhere (pp. 951, 964). Its rôle in regulating adrenal cortex activity is considered below.

(1) Effects of Hypophysectomy.—This is followed by atrophy of the adrenal cortex. As the conversion of the precursors into active hormone and its secretion into the blood is greatly depressed, the cells remain well filled with cholesterol esters and ascorbic acid. The secretory activity of the cortex does not wholly cease but continues at a low level, which is adequate for *optimal* environmental conditions, but not for conditions of stress. The animal does not die of adrenal insufficiency. Similarly, patients with Simmonds' disease (in which the anterior pituitary is severely damaged) do not show signs of Addison's disease.

(2) Injection of ACTH.—This profoundly affects the adrenal cortex. A single injection causes rapid secretion of corticoids with exhaustion of the store of precursors: the sudanophil granules disappear and the concentration

of cholesterol esters and ascorbic acid falls rapidly (Fig. 601). Subsequently the precursors are fully restored.

The gland cholesterol is derived partly from blood cholesterol; after an injection of ACTH the level of blood cholesterol ester temporarily falls. The gland cholesterol may also be synthesized locally from "acetic acid fragments" (p. 799).

If the injections of ACTH are continued over a period of days, the rate of replacement of the precursors rises and keeps pace with the rate of secretion. The gland becomes considerably enlarged (Fig. 601) and well packed with sudanophil granules which also extend into the deeper part of the cortex (zona reticularis); the concentration of cholesterol esters and ascorbic acid remains normal. Changes occur in the organism resembling those produced by injection of adrenal cortex extracts (p. 964). ACTH has *no* physiological effects in the absence of the adrenal cortex; all its actions are thus mediated by the cortex.

(3) Over-secretion of ACTH produces a syndrome of hypercortism (*Cushing's syndrome*, p. 965).

(4) Response to Stress.—Many adverse conditions (*states of stress*) internal and external cause a secretion of adrenal corticoid. These include the administration of anæsthetics like chloroform or ether, injection of insulin producing hypoglycæmia, diphtheria or tetanus toxin, infectious diseases, histamine, hæmorrhage, traumatic shock, external heat or cold, anoxia and muscular exercise. The histological changes in the cortex depend on whether the stress is mild or severe, develops quickly or gradually, is short-lasting or prolonged. All the changes to be described depend on the integrity of the hypophysis, *i.e.* they are *mediated by ACTH*. Several types of reaction may be recognized.[1]

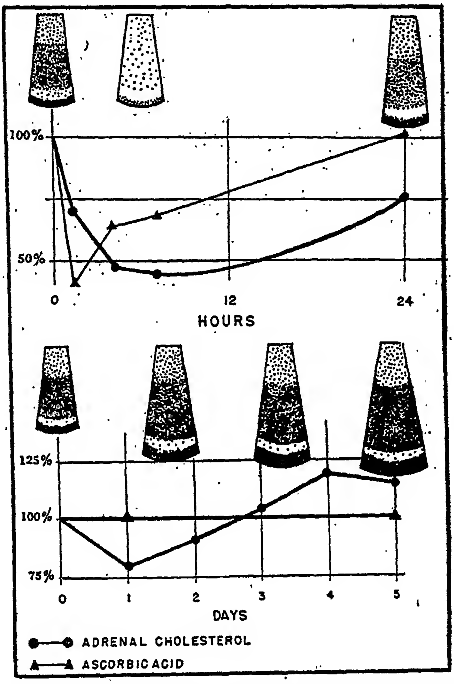


Fig. 602.—Response of Adrenal Cortex to Stress (Sayers and Sayers, *Recent Progress Hormone Research*, 1948, *2*, 81.)

Upper Record: Type I response, to sudden short-lasting stress.
Lower Record: Type II response, to gradual persisten stress.
For description see text.

(i) *Type I:* the stress is sudden and short-lasting (*e.g.* a short bout of exercise, a non-fatal hæmorrhage of one hour's duration, very short-acting drugs). There is an immediate discharge of hormone into the blood. The

[1] Sayers and Sayers, *Recent Progress Hormone Research*, 1948, *2*, 81.

sudanophil granules disappear; the cholesterol and ascorbic acid content falls. At this stage the rate of precursor breakdown exceeds that of re-formation. Later full recovery occurs and the gland becomes slightly enlarged (Fig. 602).

(ii) *Type II:* the stress develops very gradually and acts for a long time (*e.g.* pregnancy, external cold or heat). The gland progressively enlarges as the cells increased in number. The secretion of the hormone is correspondingly increased but the precursors are re-formed as rapidly as they are broken down (Fig. 602).

(iii) *Type III:* overwhelming stress ending in death (*e.g.* fatal acute infection or hæmorrhagic shock). The gland is completely depleted of its store of precursor; some hypertrophy occurs. A massive secretion of hormone is presumably taking place (Fig. 603).

(iv) *Type IV :* recovery from severe stress (*e.g.* severe chronic anoxia). Following the initial period of exhaustion (as in Type III) the hypertrophied gland becomes refilled to a greater extent than normal with precursors (Fig. 603).

(5) Mechanisms of Response.—The question arises, how does the stress stimulus cause a secretion of ACTH and consequently a secretion of corticoids? The following possibilities must be considered:

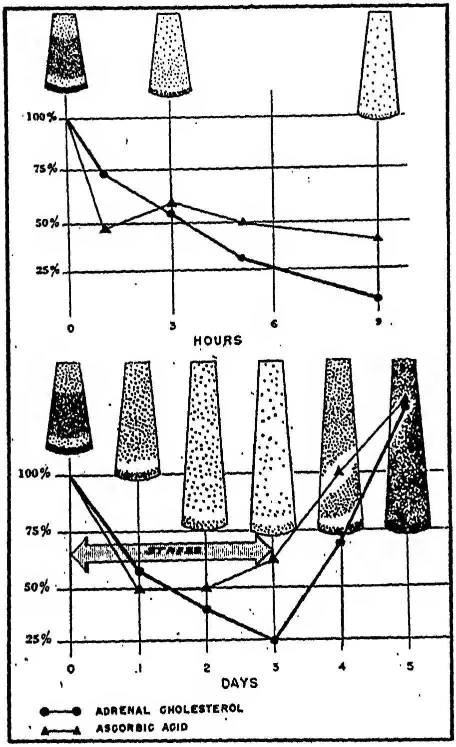


Fig. 603.—Response of Adrenal Cortex to Stress. (Sayers and Sayers, *Recent Progress Hormone Research*, 1948, *2*, 81.)

Upper Record: Type III response, to acute fatal stress.
Lower Record: Type IV response, to severe stress with recovery.
For description see text.

(i) *Nervous Mechanism.*—Stimulation of the hypothalamus causes a secretion of ACTH. The link between the hypothalamus and the anterior pituitary is not a direct nervous one; but a chemical mediator ("neurohumour") is formed in the hypothalamus which is carried in the local blood supply to the anterior pituitary as explained on p. 931. The state of stress may stimulate appropriate afferents or certain parts of the central nervous system and thus lead to hypothalamic activity which causes secretion of ACTH.

(ii) *Humoral Mechanism.*[1]—The secretion of ACTH might be regulated by the blood level of the hormone formed by the gland on which it acts

[1] Sayers, *Physiol. Rev.*, 1950, *30*, 241.

(*i.e.* the blood level of adrenal corticoid); a rise of blood corticoid level might inhibit ACTH secretion by a direct action on the pituitary, while a fall of blood corticoid might increase ACTH secretion. Such a "pituitary-

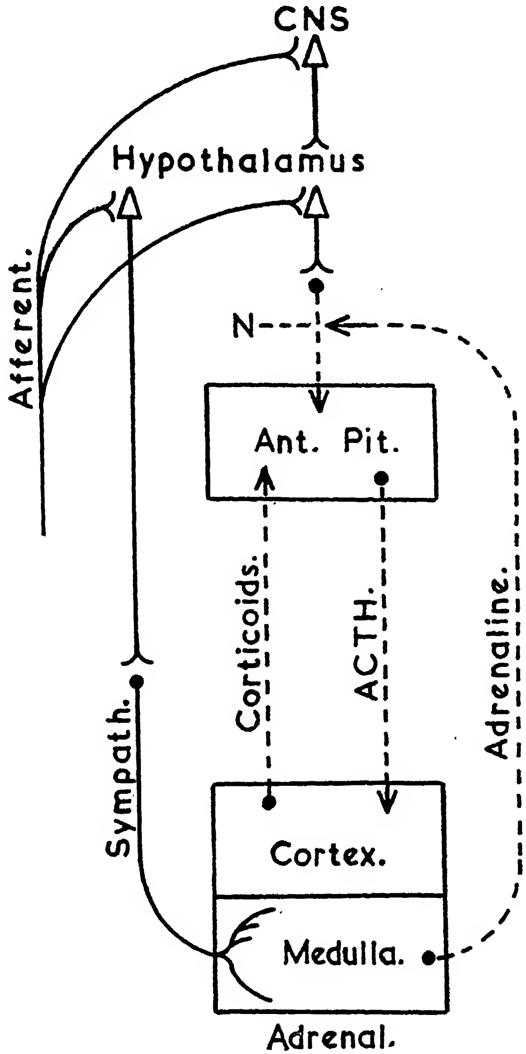


FIG. 604.—Inter-relationship of Hypothalamus, Anterior Pituitary, and Adrenal Gland.

adrenal cortex axis" would maintain a steady blood level of corticoid in the same way as the "pituitary-thyroid axis" maintains the blood thyroxine level. It is supposed that states of stress cause increased *utilization* of

corticoids in the affected tissues and consequently a fall in their blood level leading secondarily to a compensatory secretion of ACTH and an increased discharge of corticoids from the adrenal cortex. It is worth mentioning by way of caution that almost nothing is known about the fate of the corticoids in the tissues on which they act (so-called "target cells"). The following types of experiment support the hypothesis set out above.

(*a*) An animal is subjected to stress and simultaneously an active adrenal cortex extract is injected. If the dose has been accurately judged there are no histological or chemical changes in the adrenal cortex. The injected corticoids have thus prevented the fall of blood corticoid which is normally

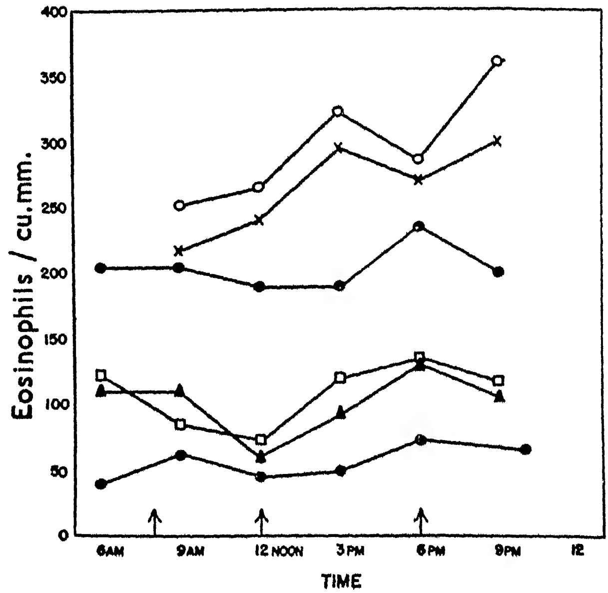


FIG. 605.—Diurnal Variations in Eosinophil Count in Normal Subjects. (Thorn *et al.*, *J. clin. Endocrin.*, 1950, *10*, 198.)

produced by stress; in consequence ACTH is not secreted and the cortex does not respond. Such experiments also enable the amount of corticoid secreted in various states of stress to be roughly measured.

(*b*) Injection of adrenal cortex extracts into an intact animal causes adrenal atrophy; the rise of blood corticoid has inhibited ACTH secretion. Conversely partial removal of the adrenals, by reducing corticoid secretion, stimulates the pituitary and the excess ACTH released causes hypertrophy of the remaining cortical tissue.

(iii) *Adrenaline as Mediator.*—Injection of adrenaline causes secretion of ACTH, probably owing to a direct action on the pituitary gland. Any condition, therefore, which causes the secretion of an adequate dose of adrenaline may lead secondarily to a secretion of corticoids (p. 728).

The mechanisms which are probably concerned in regulating the activity of the adrenal cortex are summarized in Fig. 604.

Clinical Indices of Corticoid Secretion.[1]—(1) BLOOD EOSINOPHIL COUNT.—The normal eosinophil count in the blood in man is 50–500 per c.mm.; the fluctuations in any individual during the 24 hours are, however, comparatively small (Fig. 605), though the count tends to rise in the afternoon.

If adequate control counts have been made, a change of 40% or more produced experimentally is significant.

(i) Injection of *adrenal cortex extracts* markedly lowers the eosinophil count; how the change is produced is unknown.

(ii) A similar result is obtained by injecting *ACTH* which acts solely by releasing adrenal corticoids (Fig. 606). ACTH is still effective in rats after

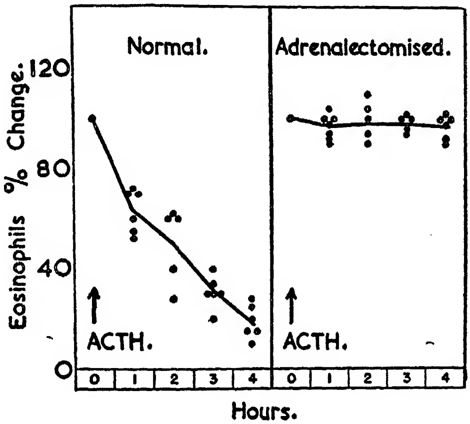


FIG. 606.—Effect of ACTH on Eosinophil Count in Normal and Adrenalectomized Rats. (Thorn *et al.*, *J. clin. Endocrin.*, 1950, *10*, 203.)

Ordinate: % change in eosinophil count.
120, 80, 40=120%, 80%, 40% of control level (100=control level).

destruction of the adrenal medulla; it produces no change in the eosinophils after complete adrenalectomy (Fig. 606).

(iii) Intravenous injection of *adrenaline* stimulates a secretion of ACTH and thus of adrenal corticoid and so lowers the eosinophil count. Adrenaline fails to produce a response after hypophysectomy or after adrenalectomy. (Fig. 607).

These observations are of great interest clinically and experimentally. If suitable controls have been carried out an experimentally induced fall of the eosinophil count is proof of the occurrence of corticoid secretion; the amount of corticoid thus secreted can be assessed by determining the dose which produces an equivalent change in the eosinophils. Using this technique, it can be confirmed in man that many types of stress and disease cause corticoid secretion, *e.g.* surgical operations, trauma, infections, cardiac infarction. The eosinophil response can also be used to assess the functional state of the anterior pituitary and the adrenal cortex in patients.

[1] Hills *et al.*, *Blood*, 1948, *3*, 755. Thorn *et al.*, *J. clin. Endocrin.*, 1950, *10*, 187.

Clinical Procedures.—(i) *Adrenaline Test.*—Infuse 0·2 mg. of adrenaline in 200 c.c. of saline intravenously in one hour. In normals the eosinophil count 4 hours later falls on an average by 60%. In patients with severe hypopituitarism or Addison's disease no change occurs (Fig. 608).

Total White Cell Response to Adrenaline.—With the dose indicated the total white cell count after 4 hours is increased owing to a rise in the neutrophils; there is an irregular rise in the lymphocytes followed by a fall.

(ii) *ACTH Test.*—Inject 25 mg. of ACTH intramuscularly. In normals the average fall in the eosinophils 4 hours later is 75%. In Addison's disease there is no response. In hypopituitarism the response is smaller than normal owing to the atrophic condition of the adrenal cortex (Fig. 608).

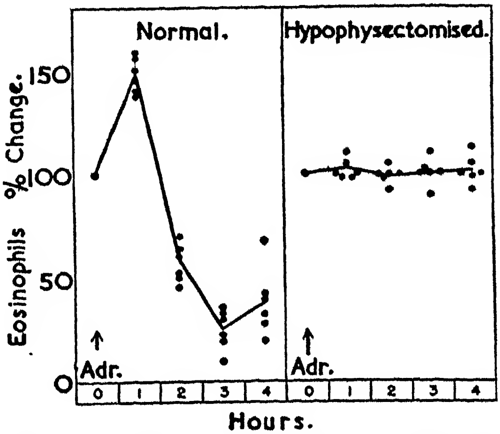


Fig. 607.—Effect of Adrenaline on Eosinophil Count in Normal and Hypophysectomized Rats. (Thorn *et al.*, *J. clin. Endocrin.*, 1950, *10*, 206.)

Ordinate: % change in eosinophil count.
100 = control level.
150, 50 = 150%, 50% of control level.

(iii) *Corticoid Test.*—A positive response with this test proves that that unknown peripheral mechanisms which lower the eosinophil count are reactive. The adrenal amorphous fraction, cortisone and compound F lower the eosinophil count in man; this result leaves them on the "short list" as claimants to the title of the natural cortex hormone. Desoxycorticosterone is ineffective.

Cushing's Syndrome.—In this condition, which is associated with hypercorticalism, the eosinophil count, as expected, is low; the finding is of diagnostic value (p. 967).

(2) Na^+ Content of Sweat as Index of Corticoid Secretion.—As explained on p. 945 the adrenal corticoids decrease the Na^+ content of sweat (Fig. 611). Changes in the composition of the sweat thus reflect alterations in the rate of corticoid secretion.

(3) Excretion of Neutral 17-ketosteroids and Adrenal Corticoids in the Urine.[1]—(i) The 17-*ketosteroids* (as their name suggests) are

[1] Mason and Engstrom, *Physiol. Rev.*, 1950, *30*, 321.

steroids with an =O attached to the C in the 17 position (as for example in androsterone, Fig. 610). These ketosteroids are end-products of corticoid metabolism in both sexes, but in the male they are also produced to a small extent by the testis. [Though œstrone, which is excreted in the urine in

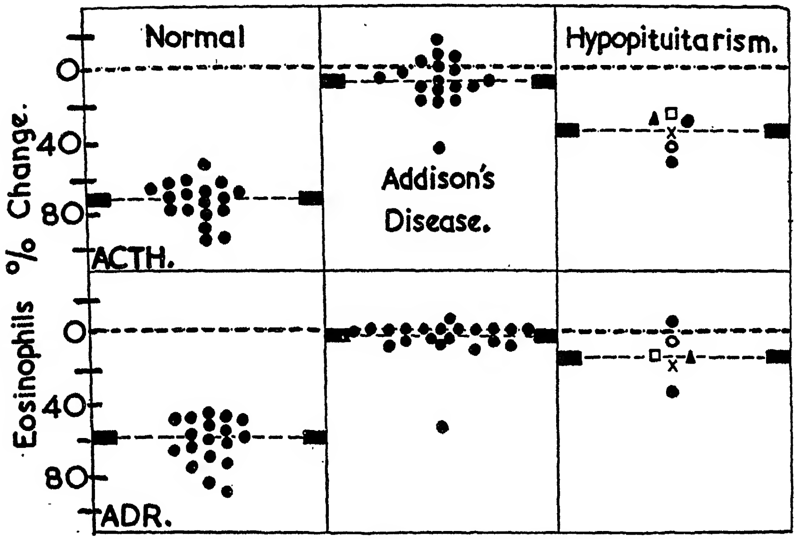


FIG. 608.—Effect of Adrenaline and ACTH on Eosinophil Count in Normal Subjects, in severe Addison's Disease, and in Hypopituitarism. (Thorn *et al.*, *J. clin. Endocrin.*, 1950, *10*, 221.)

0=Control level. 40, 80=40%, 80% *decrease* compared with control level.
Upper Records: % change in eosinophil count 4 hours after injecting 25 mg. of ACTH intramuscularly.
Lower Records: % change in eosinophil count 4 hours after infusing 0·2 mg. of adrenaline intravenously.
The short black bars joined by a dotted line indicate the mean changes.

women, is a 17-ketosteroid, it is a phenolic (*acidic*) compound and is therefore *not* estimated in the methods used for *neutral* 17-ketosteroids.] The normal range of urinary output[1] of neutral 17-ketosteroids (in mg. per 24 hr.) is 11–21 in men, and less in women, 4·5–19·5. In conditions of severe cortical insufficiency (*e.g.* Addison's disease, Simmonds' disease with cortical atrophy (p. 931)), the output in the urine is reduced to subnormal levels, and is not increased by an injection of ACTH as is the case when the gland is normal. The Table below illustrates these points:

URINARY NEUTRAL 17-KETOSTEROIDS IN MG. PER 24 HR.

	Normal.	Addison's Disease.
Control	9·0	3·0
Inject ACTH (40 mg. daily for 6 days)	17·0	2·9

[1] Butt *et al.*, *Lancet* 1950, ii 894.

The ketosteroid output may be increased when there is cortical hyperplasia, but the increase is usually striking when a *secreting tumour* (especially carcinoma) is present. Fig. 609 shows that in a group of cases with simple hyperplasia the neutral 17-ketosteroid excretion rarely rose above 50 mg. In the cases with carcinoma the values usually exceeded 50 mg. and in some instances were much higher (*e.g.* 250 mg. (or even 2000 mg.)). The excretion of neutral 17-ketosteroids is therefore increased in Cushing's syndrome and in the adrenogenital syndrome, but it is not a sensitive index of cortical function and its determination is of little value in doubtful cases.

(ii) *Urinary Adrenal Corticoids* (*so-called* 11-*oxysteroids*).—These are substances which in their biological action and chemical properties resemble the natural adrenal corticoids. The amount in the urine may be assayed by *biological* methods, *i.e.* it is tested for corticoid physiological activity (*e.g.* power to increase liver glycogen or to prolong life in adrenalectomized animals), or it may be determined by *chemical* methods. Sayers suggests that the terms " biocorticoid " and " chemocorticoid " might be used to designate respectively the substances which are thus measured. The output of the urinary corticoids correlates well with cortical activity. The normal output (per 24 hr.) is equivalent in biological activity to 0·1 mg. of cortisone. Corticoid excretion is decreased in Addison's disease and in Simmonds' disease, and is increased in clinical states of hypercortism and after injection of ACTH if the adrenal cortex is not subnormal.

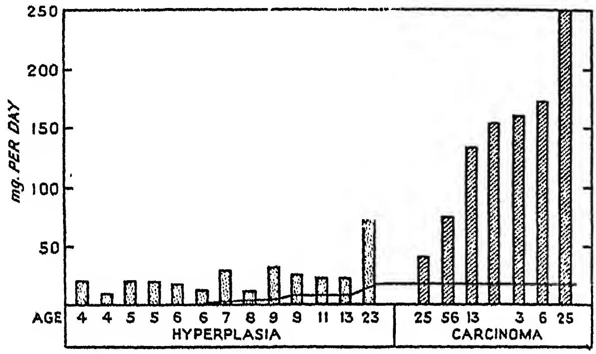


FIG. 609.—Neutral 17-ketosteroids in Urine in Hyperplasia and Carcinoma of Adrenal Cortex. (After Talbot *et al.*, *J. clin. Investig.*, 1942, *21*, 560.

Ordinate: Excretion of 17-ketosteroids in urine in mg. per day.
Continuous line: Upper normal limit of 17-ketosteroid excretion in children (left) and adults (right).
For description see text.

Both cortisone and compound F have been isolated from the urine in normal subjects and in cases of Cushing's syndrome.

Adrenal Cortex Insufficiency. Addison's Disease.[1]—Addison's disease is the classical clinical syndrome of adrenal cortex insufficiency. The whole gland is usually affected, cortex as well as medulla; but the clinical manifestations are due almost entirely to lack of cortical secretion. The

[1] Gaunt *et al.*, *Physiol. Rev.*, 1949, *29*, 281. Thorn *et al.*, *Adrenal Insufficiency*, Oxford, 1949. Sorkin, *Medicine*, 1949, *28*, 371.

lesions most commonly found are simple atrophy and tuberculous infection. The main clinical features are: muscular weakness, low blood pressure, pigmentation, complex changes in fluid and ionic balance and in general metabolism, and terminally, renal and circulatory failure.

The detailed account that follows is amplified where necessary by reference to findings in adrenalectomized animals.

(1) ALTERATION IN IONIC AND FLUID BALANCE.—(i) There are characteristic changes in the ionic composition of the plasma: (*a*) There is generally a fall in plasma Na^+ to 135 m.Eq./L or less (average normal, 142 m.Eq./L). Thus in one series of patients the plasma Na^+ value at some stage of the disease was between 115 and 130 m.Eq./L.

(*b*) Less regularly there is a rise of plasma K^+; values exceeding 6 m.Eq./L (23·5 mg–%) are pathological (average normal, 5 m.Eq./L=19·5 mg–%).

(*c*) A fall of plasma Cl′ (average normal, 103 m.Eq./L) and of bicarbonate (alkali reserve), (average normal 60 c.c. of CO_2/100 c.c. of plasma) is common. The plasma changes are probably brought about as follows:

(*α*) There is a primary derangement of the mechanisms concerned with reabsorption of Na^+ and K^+ by the renal tubular epithelium. There is decreased reabsorption of Na^+ and therefore increased loss in the urine: plasma Na^+ consequently falls. On the other hand there is increased reabsorption of K^+ and, therefore, decreased loss in the urine; plasma K^+ consequently rises. As the fall in plasma Na^+ exceeds the rise in plasma K^+ the total concentration of plasma cations decreases.

(*β*) The treatment of Cl′ in the tubules is more variable and does not necessarily follow that of Na^+. If the reabsorption of Cl′ is depressed to the same extent as that of Na^+, there is an equivalent fall of Na^+ and Cl′ in the plasma and the blood reaction is unaltered. If there is a relatively greater loss of Na^+ in the urine, some of the $NaHCO_3$ of the plasma loses its Na^+; there is a corresponding fall in the alkali reserve and a condition of metabolic acidosis develops.

(ii) The consequent changes in the body fluids resemble those found in simple Na^+ or NaCl deficiency. Ionic equilibrium is rapidly attained between the plasma and the interstitial fluid. Owing to the functional "impermeability" of the tissue cell membranes to Na^+, the intracellular crystalloid osmotic pressure initially exceeds that in the extracellular fluid. There is, therefore, a flow of water from the extracellular into the intracellular fluid. The *plasma volume is decreased*: the plasma protein concentration, the red cell count and the hæmoglobin concentration are all increased, *i.e.* there is *hæmoconcentration* which is secondary to the anhydræmia.

(iii) The changes in the blood lead to *circulatory failure* with decreased blood pressure and diminished blood flow to the organs.

(iv) The *kidney* is adversely affected in two ways: (*a*) Its functional activity is very dependent on its blood supply; there is a decrease in renal blood flow and probably also in glomerular pressure which lead to decreased glomerular filtration, decreased flow of urine (or even anuria) and decreased excretion of nitrogenous waste products.

(*b*) As is explained on p. 64 decreased plasma Na^+ depresses renal activity in various ways: there is a further decrease in glomerular filtration,

a *delayed* or *inadequate response to water drinking* (which thus causes an excessive accumulation of body fluid) and a diminished excretion of urea (cf. p. 65). The kidney also fails to make ammonia to compensate for the acidæmia.

(v) The fatal outcome is due to the combined effects of circulatory and renal failure, including the changes in the blood composition.

(vi) The lack of ionic balance is probably responsible for certain *metabolic disturbances* such as impaired absorption of carbohydrate and fat from the small intestine, decreased metabolic rate and loss of appetite.[1]

(vii) All the changes enumerated above (i–vi) can be rectified by giving a diet rich in Na^+ (*e.g.* 10 g. of NaCl daily) and poor in K^+. Similar improvement can be produced without necessarily changing the diet by administration of: (*a*) the *amorphous fraction* of the adrenal cortex; or (*b*) *desoxycorticosterone*; or (*c*) *cortisone*. Less hormone is needed if the diet is suitably adjusted. The symptoms of the disease are aggravated or a dangerous crisis may be precipitated by restricting the intake of sodium chloride.

Clinical Tests.—Two tests of adrenal insufficiency are based on the disturbances described above.

(*a*) The response to water drinking is tested (Robinson's test); the resulting diuresis develops more gradually in patients than in normal subjects.

(*b*) The patient's ability to maintain the plasma Na^+ concentration within normal limits, when the Na^+ intake is drastically reduced, is tested. The Na^+ intake is reduced to less than 1 g. daily and hormone treatment (if any was being given) is stopped. In a case of Addison's disease, even if the plasma Na^+ was previously normal it will fall below 135 m.Eq./L within a few days, and the characteristic symptoms of the disease appear or become aggravated. The patient must be carefully watched as intensive treatment must be started should the clinical state deteriorate.

(2) Changes in General (Carbohydrate, Protein, and Fat) Metabolism.—These changes occur in adrenalectomized animals and in patients even when the diet is rich in Na^+ and poor in K^+. The results of adrenalectomy resemble in many respects those of hypophysectomy. It should be remembered that the pituitary diabetogenic factor is partly ACTH and that, in the main, the adrenal cortex is an antagonist to insulin.

(i) *Blood Sugar.*—(*a*) Owing to cortical deficiency, tissue protein is not broken down to amino-acids to be subsequently used for glucogenesis; similarly hepatic glucogenesis from amino-acids is depressed; tissue uptake of glucose is perhaps increased. As a result the blood sugar falls slightly even in the well fed adrenalectomized animal. But if the animal is starved the failure of glucogenesis causes a severe and possibly fatal hypoglycæmia. Some degree of hypoglycæmia is common in patients with Addison's disease and produces its characteristic symptoms.

(*b*) There is excessive insulin sensitivity; *i.e.* a small dose of insulin produces a marked fall of blood sugar, followed by a slow recovery.

(*c*) Liver glycogen stores are low.

(ii) *Muscle Power.*—The muscles are weak and quickly become fatigued

[1] Some patients may display grave symptoms of adrenal insufficiency although the ionic balance is normal.

on stimulation. The cause of the weakness is not known with certainty. It may be due to the abnormal ionic pattern in the body fluids; it is usually benefited by administering Na^+ salts and corticoids which raise plasma Na^+. In a few patients, power is not increased by these measures. Experimentally, the power of adrenalectomized animals is improved by infusions of glucose or injection of "glucocorticoids," suggesting that a defect of carbohydrate and related metabolism is a responsible factor.

(3) BLOOD PRESSURE.—The blood pressure is commonly low; values of 90 mm. systolic, 65 mm. diastolic are typical. The patient may become dizzy and faint when passively tilted from the horizontal to the erect position. If a state of "crisis" develops still lower values occur. The blood pressure is raised to normal and even to abnormally high levels by Na^+ salts and by corticoids like desoxycorticosterone.

(4) PIGMENTATION.—The pigment which is deposited in the skin is melanin, derived from tyrosine by the action of the enzyme tyrosinase. The cause of the pigmentation is unknown; it does not occur in adrenalectomized animals. As tyrosine is also a precursor of adrenaline the pigmentation has been accounted for as follows: the destruction of the adrenal medulla leads to the accumulation of tyrosine which would normally be converted into adrenaline; this unused tyrosine is transformed into melanin. This explanation is inaccurate as pigmentation may occur in patients in whom the medulla is intact. It has been suggested that the enzyme reactions leading to melanin formation may be under hormonal control; the pigmentation is, however, little affected by administration of corticoids.[1]

(5) NAUSEA AND VOMITING are common; they are attributed to loss of Na^+ and to acidæmia. The similarity between many of the findings in Addison's disease and sodium chloride deficiency must be emphasized (p. 65).

(6) CRISIS.—Acute cortical insufficiency produces the so-called *crisis*. There is vomiting, apathy, drowsiness and signs of dehydration; the blood pressure falls to a very low level; the terminal signs are delirium, coma, and evidence of grave renal failure.

(7) TERMINAL CHANGES IN ADRENALECTOMIZED ANIMALS.—Towards the end many metabolic processes are depressed; the rate of deamination of amino-acids, the rate of formation of ketone bodies (in the liver) and their utilization by the tissues, phosphorylation processes in general, and the level of tissue O_2 consumption are all decreased; body temperature finally falls. These changes are apparently secondary to the ionic disturbances as they clear up when the latter are corrected.

(8) URINE.—In adrenal cortex deficiency the urinary excretion of the neutral 17-ketosteroids is reduced. The excretion of the corticoids in the urine may cease (p. 954).

Active Principles of Adrenal Cortex. Corticoids.[2]—It is customary to divide the physiologically active substances which can be extracted from the adrenal cortex into two categories: (i) *crystalline* components; (ii) the *amorphous* fraction.

(1) CRYSTALLINE COMPOUNDS.—About 50 have been isolated from the adrenal cortex or synthesized but of these only a few possess corticoid

[1] Lerner and Fitzpatrick (Melanin formation), *Physiol. Rev.*, 1950, *30*, 91.

[2] Heard, in *The Hormones*, N.Y., 1948, *1*, 550. Hechter, *Recent Progress Hormone Research*, 1951, *6*, 215. The elementary structure of the steroids is described on p. 1074.

actions, *i.e.* actions resembling in some respects those of adrenal cortex extracts. These steroids can be divided into three main groups (Fig. 610).

(i) *Oxy*corticoids with an =O or —OH grouping attached to the C in the 11 positions, *e.g.* corticosterone, cortisone (Kendall's compound E), 17–hydroxycorticosterone (compound F), and 11–dehydrocorticosterone. These steroids correct the abnormalities of *general* metabolism (*i.e.* those of carbohydrate, protein, and fat metabolism) in adrenalectomized animals. They are generally called *glucocorticoids* because of their easily observed effects on carbohydrate metabolism.

(ii) Steroids without an =O or —OH in the 11 position, *i.e. desoxy*-corticoids. They have no direct effect on general metabolism but they correct the abnormalities of ionic equilibrium in the adrenalectomized animal. They are therefore called *mineralocorticoids*. The best known are desoxycorticosterone (deoxycortone) and desoxycortisone (compound S). Desoxycorticosterone acetate is commonly abbreviated to DOCA or DCA.

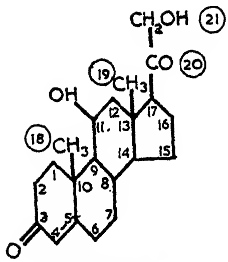
The subdivision into so-called gluco- and mineralocorticoids is however only of limited value. The mineralocorticoids, because they restore the normal ionic balance, also remedy the carbohydrate and other disturbances which are secondary to the ionic imbalance. Even more significant is the fact that the oxycorticoids cortisone and compound F, in suitable dosage, can directly influence both the ionic imbalance and the derangements of general metabolism. Cortisone is thus as much a mineralo- as a gluco-corticoid. The term glucocorticoid is unsatisfactory for another reason: the metabolism of the foodstuffs is closely inter-related; the metabolic disturbances in Addison's disease which are corrected by the active corticoids involve proteins and fats as well as carbohydrates. It would be more logical to designate the so-called glucocorticoids as metabolocorticoids.

(iii) Sex hormones like progesterone, œstrogens and androgens; the adrenal androgens resemble *androsterone* in having an =O in the 17 position, *i.e.* they are neutral 17–ketosteroids. One androgen isolated from the adrenal is called *adrenosterone*. Adrenal cortex extracts may contain a substance called cortilactin which influences the development of the breast. Neutral 17–ketosteroids of adrenal origin are normally excreted in the urine (p. 952).

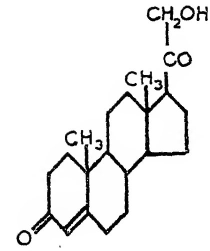
(2) Amorphous Fraction.—This is of undetermined composition. It is a disconcerting fact that this material represents 90% of the activity of crude adrenal cortex extracts as tested by their power to maintain life and to regulate ionic balance in adrenalectomized animals; the chemically identified steroids of the adrenal cortex represent only 10% of the curative activity of crude adrenal extracts.

Natural Adrenal Cortex Hormone.—There is still uncertainty about the identity of the steroid or steroids or other active principles secreted by the adrenal cortex. The following observations throw some light on the subject:

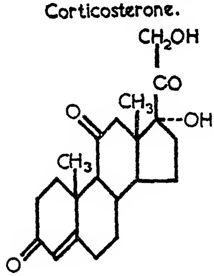
(i) The only physiological action of ACTH is to cause the adrenal cortex to secrete its hormones. The results of injecting ACTH in normal subjects can be compared with the effects produced by various corticoids. The corticoid which most closely reproduces the action of ACTH would have a claim to be regarded as *the* natural hormone or at least as *one* of the natural hormones. As explained later both cortisone and compound F satisfy this criterion.



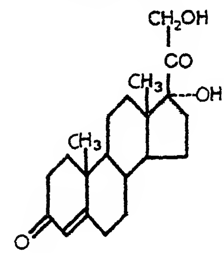
Corticosterone.



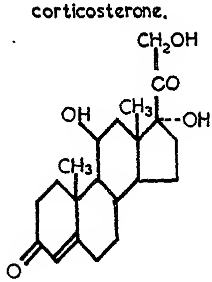
Desoxycorticosterone.
Deoxycortone.



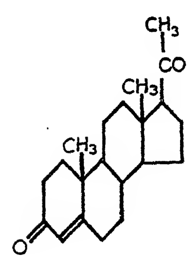
Cortisone.
Kendall's Compound E.
11-dehydro-17-hydroxy-corticosterone.



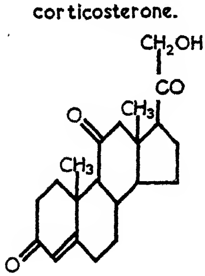
Desoxycortisone.
Reichstein's Compound S.
11-desoxy-17-hydroxy-corticosterone.



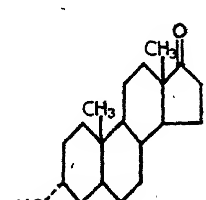
Compound F.
17-hydroxy-corticosterone.



Progesterone.



11-dehydrocorticosterone.
[=Cortisone minus OH at position 17]



Androsterone.

FIG. 610.—Structure of Adrenal Corticoids.

in blood pressure and an increase in muscular strength. DOCA represents only a partial form of replacement therapy as it does not correct the abnormalities of general metabolism; it does not prevent hypoglycæmia. It has no effect on the eosinophil count. It has, however, considerably prolonged the expectation of life in patients with Addison's disease.

Toxic Effects of Deoxycortone.—(i) Large doses produce toxic symptoms in *patients*. There is excessive retention of Na^+ and Cl' and secondarily of water, leading to increased plasma volume, extensive œdema, hypertension,[1] and cardiac embarrassment; the stores of K^+ are depleted and prostration may occur. The plasma Na^+ may exceed 140 m.Eq./L and the plasma K^+ may fall below 3 m.Eq./L. The symptoms are aggravated by a high NaCl intake. When deoxycortone is used in the treatment of Addison's disease it is best combined (for safety) with a reduced Na^+ and high K^+ intake.

(ii) In *experimental animals*, overdosage with deoxycortone is alleged to produce (in addition to hypertension) pathological changes in many organs;[2] these include malignant nephrosclerosis, necrotic changes in the arteries resembling those found in periarteritis nodosa, the development of Aschoff bodies in the heart like those of rheumatic fever, and joint lesions like those of rheumatoid arthritis. The mineralocorticoid desoxycortisone (compound S) produces similar lesions experimentally. It is alleged that these structural changes can be prevented by administering suitable doses of the oxycorticoid cortisone.[3]

(3) **Cortisone.**[4]—Very thorough clinical and experimental studies have been carried out with this steroid. In doses of 10–20 mg. daily it maintains health in cases of *Addison's disease:* it is also beneficial in cases of Simmonds' disease (p. 930). Given in daily doses of 50–100 mg. (or less) it has proved beneficial in an extraordinary range of diseases *in none of which there is any evidence of disturbed secretion of adrenal corticoids.* It is as effective by mouth as by intramuscular injection.

Injection of ACTH produces similar therapeutic effects in these diseases if the adrenal cortex is intact, presumably by releasing a cortisone-like substance.

The results of detailed studies of the action of cortisone are summarized below:

(i) General Metabolism.—(*a*) Like the other oxycorticoids, cortisone *opposes* in some respects the action of insulin. Thus it raises the blood sugar, and increases neoglucogenesis. Glycosuria may develop; this is due partly to the hyperglycæmia and partly perhaps to a phloridzin-like action on the renal tubules paralyzing the reabsorption of glucose. It aggravates all the

[1] DOCA may have a direct constrictor action on blood vessels.

[2] Selye, *Brit. med. J.*, 1950, i, 203.

[3] *Adrenal Cortex and "Adaptation Syndrome."* According to Selye, the bodily responses to "stress" comprise what he has termed the "*adaptation syndrome*" which is divided arbitrarily into three stages: (i) an initial *alarm reaction*; (ii) *adaptation* proper, when the body's compensatory reactions to stress develop; and (iii) a phase of *exhaustion* which occurs if exposure to stress is excessive. As explained on p. 947, adrenal corticoid secretion occurs in many states of stress. (Selye, *J. clin. Endocrin.*, 1946, *6*, 118.)

[4] Carlisle, *Brit. med. J.*, 1950, ii, 590. Hench *et al.*, *Proc. Staff Mayo Clinic*, 1949, *24*, 181. Symposium, *ibid.*, 1950, *25*, 474–502. Thorn and Forsham, *Recent Progress Hormone Research*, 1949, *4*, 229. Sprague *et al.*, *J. clin. Endocrin.*, 1950, *10*, 289. Sprague *et al.*, *Arch. int. Med.*, 1950, *85*, 199. Hench *et al.*, *Arch. int. Med.*, 1950, *85*, 545. Ingle, *J. clin. Endocrin.*, 1950, *10*, 1312. Sprague *et al.*, *Recent Prog. Hormone Res.*, 1951, *6*, 315.

signs in diabetic patients. On stopping cortisone treatment transient hypoglycæmia may develop ; this result is attributed to cortisone depressing ACTH secretion and consequently the normal activity of the adrenal cortex. Again like the other oxycorticoids (but in this respect acting *like* insulin) cortisone increases the deposition of glycogen in the liver.

(*b*) It increases the breakdown of tissue proteins ; the released amino-acids in part give rise to blood glucose as mentioned above.

(*c*) The effect on total nitrogenous excretion in the urine is variable ; it is often increased. Urinary creatinine is unaffected ; the output of creatine (if creatinuria was previously present) and of uric acid is increased. The increase in the ratio of uric acid/creatinine or of creatine/creatinine is a convenient index of the urinary changes. In patients with gout, uric acid excretion is strikingly increased by administration of cortisone ; this effect is in part due to decreased reabsorption of uric acid from the renal tubules into the blood.

(*d*) Cortisone may increase the mobilization and utilization of fat.

(*e*) Calcium (and phosphate) excretion may be increased, resulting in osteoporosis.

(ii) Body Fluids and Ionic Balance.—Like the desoxycorticoids, cortisone corrects the ionic imbalance and the other body fluid disturbances which are found in Addison's disease. There is thus retention of Na^+ and water (and usually of Cl′) and increased excretion of K^+.

In *toxic* doses cortisone produces the same undesirable effects as deoxycortone. There is excessive retention of Na^+ and water ; if this condition is not dealt with by a natural diuresis, there is œdema (peripheral or pulmonary), ascites, congestive heart failure and hypertension. (Treatment consists of reducing the Na^+ intake and administering diuretics.)

If Cl′ is not retained in equivalent amounts to Na^+, the excess Na^+ unites with HCO_3' to form sodium bicarbonate, thus increasing the plasma alkali reserve and producing alkalosis.

There may be excessive excretion of K^+ in the urine. The resulting low serum K^+ causes muscular weakness (cf. periodic paralysis, p. 519), arterial hypotension and changes in the electrocardiogram (Fig. 615). The symptoms are relieved by administering K^+ salts.

(iii) Muscle Power is restored by cortisone in cases of Addison's disease and in adrenalectomized animals.

(iv) Central Nervous System.—The patient may feel well and cheerful ; or he may sleep badly and become restless, excitable, or even maniacal. Some patients become depressed. The frequency of the *alpha* waves in the electroencephalogram is increased.

(v) Blood.[1]—(*a*) As a result of cortisone treatment the lymphoid tissues, including the thymus, atrophy. The number of *circulating* blood lymphocytes is reduced. It has been suggested that there is increased destruction of lymphocytes within the lymphoid tissues.

(*b*) *The circulating eosinophil leucocytes are decreased.*

(*c*) Polycythæmia may be produced.

(vi) Large doses produce some of the characteristic features of Cushing's syndrome, including rounding of the face and, in women, growth of hair on the face, chest, and anterior abdominal wall (hirsutism), and amenorrhœa.

[1] White, *Harvey Lectures*, 1947–48, 1. Harris *et al.*, *J. Physiol.*, 1950, *111*, 328, 335.

These observations suggest that Cushing's syndrome is due to hypercortism (p. 965).

(vii) The adrenal cortex atrophies because the raised level of blood cortisone depresses the secretion of ACTH.

General Therapeutic Effects of Cortisone.—The therapeutic action of cortisone in states of adrenal cortex insufficiency, and the toxic effects described above, are all in keeping with what is known about the physiology of the adrenal cortex. Cortisone, however, has beneficial effects in the diseases set out below in the Table (from Carlisle) *in none of which is the adrenal cortex known to be at fault.* Its mode of action is so far unexplained. Of a very large number of corticoids (natural and synthetic) so far tested, only compound F has been found to have therapeutic effects like those of cortisone.[1] Even the "diene" of cortisone (*i.e.* cortisone with a second double bond joining C_6 and C_7), though it prolongs life in the adrenalectomized animal, does not possess the general therapeutic actions of cortisone.

Response of Various Diseases to Cortisone

Beneficial Effect often Dramatic	Results Encouraging but Require Further Evaluation	Transient Beneficial Effects Observed
Rheumatoid arthritis	Various allergies : Hay-fever Angioneurotic œdema Drug sensitization Serum sickness	Acute leukæmia (lymphocytic or granulocytic)
Rheumatic fever	Acute gouty arthritis	Multiple myeloma
Rheumatoid spondylitis	Ulcerative colitis	Lymphosarcoma
Still's disease	Regional enteritis	Hodgkin's disease
Psoriatic arthritis	Nephrotic syndrome	Chronic lymphatic leukæmia
Lupus erythematosus (early)	Scleroderma (early)	
Asthma (status asthmaticus)	Dermatomyositis	
Inflammatory eye diseases	Psoriasis	
Exfoliative dermatitis	Periarteritis nodosa (early)	
Pemphigus	Pulmonary granulomatosis	
	Alcoholism (?)	

An important generalization about cortisone is that it blocks allergic hypersensitivity reactions. It is consequently useful in hay fever, asthma, drug reactions, and serum sickness.

[1] Fourman *et al.*, *J. clin. Investig.*, 1950, *29*, 1462.

In patients with certain diseases, cortisone causes the tissues to behave as " though the noxious agent was not there " ; cortisone in these people acts like " an asbestos suit against fire," protecting the tissues from damage and preventing their normal responses to injurious agents, but not extinguishing the fire. As might be expected, in most instances improvement persists only for as long as treatment is continued ; a relapse occurs on stopping treatment unless in the meantime the disease has come to its natural end.[1]

Clinical Action of ACTH.[2]—ACTH has been obtained from the gland in the form of a simple protein (mol. wt. 20,000).[3] Clinically it can be employed safely and effectively. When injected into normal subjects, it

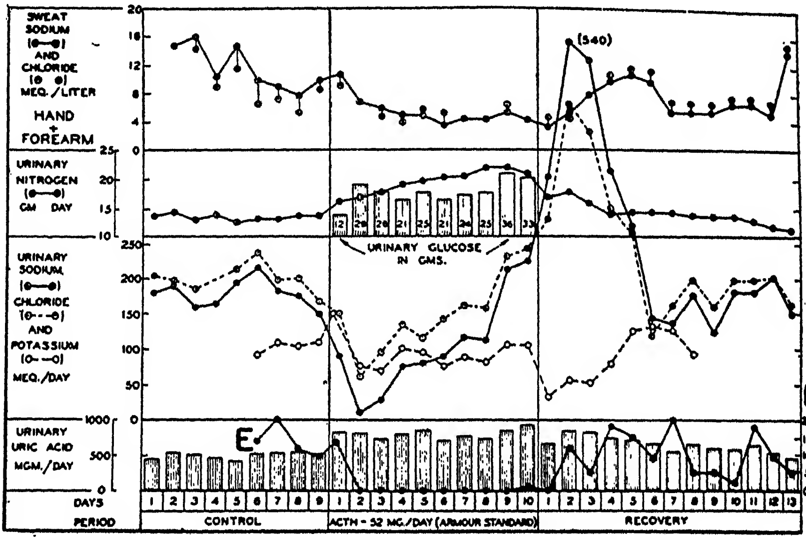


Fig. 611.—Action of ACTH in Normal Man. (Conn *et al.*, *J. clin. Endocrin.*, 1950, *10*, 12.)

produces in the main the same physiological effect as cortisone, but the adrenal cortex is enlarged (and not atrophied).

(1) The changes in the urine, sweat, and blood eosinophil count are well summarized in Fig. 611, which shows the effects of injecting 50 mg. of ACTH daily for 10 days.

(i) The Na^+ and Cl' content of the sweat is reduced.

[1] Cortisone produces alterations in the local and general responses to injury and infection. Thus it depresses the formation of granulation tissue in wounds. It diminishes the resistance of mice to experimental tuberculosis. In man it diminishes or abolishes the tuberculin reaction ; in patients it may cause *acute spread of tuberculous infection.* It prevents or reduces tissue reaction to chemical irritants (in guinea pigs).

[2] Thorn *et al.*, *J. clin. Endocrin.*, 1950, *10*, 187. Conn *et al.*, *ibid.*, p. 12. Venning *et al.*, *J. clin. Investig.*, 1950, *10*, 583.

[3] A polypeptide with a mol. wt. of about 2500 has been prepared which is about 100 times as active as this protein preparation ; a smaller active polypeptide of mol. wt. 1200 has also been prepared. It would seem that the physiological activity of the ACTH protein resides in a restricted part of the molecule.

(ii) There is a marked increase in urinary N on a constant diet, indicating that tissue protein is being broken down.

(iii) Glycosuria develops. This is due to hyperglycæmia and to decreased glucose reabsorption in the tubules.

(iv) The excretion of urinary Na^+ and Cl' is initially reduced, the former more than the latter; but after some days the effect wears off. There is a transient washing out of K^+ in the urine. On stopping the injections there is a greatly increased (compensatory) excretion of Na^+ and Cl' in the urine.

(v) Uric acid excretion rises.

(vi) The blood eosinophil count falls.

(vii) There is an increased excretion of both corticoids and neutral 17–ketosteroids in the urine.

(viii) The general clinical picture produced by ACTH (as with cortisone) resembles Cushing's syndrome.

(ix) ACTH has in general the same therapeutic effects as cortisone, but it is of course useless in Addison's disease. Its use as a diagnostic agent is considered on p. 952.

(2) ACTH is part of the *diabetogenic factor* of anterior pituitary extracts. The relative importance of ACTH and growth hormone as diabetogenic agents varies with the species. ACTH readily produces diabetic manifestations in man; pure injected growth hormone has so far been inert in man.

Clinical Manifestations of Hyperfunction of Adrenal Cortex. (Hypercortism).—Two main clinical syndromes occur:

(i) Cushing's ("Cortico-metabolic") Syndrome, in which there is excessive secretion of a corticoid which resembles cortisone in its actions.

(ii) Adrenogenital ("Cortico-sexual") Syndrome, in which there is excessive formation of sex hormones, either male (androgens) or female (œstrogen, progesterone or a lactogen ("cortilactin")).

Mixed syndromes may also occur.

Anatomical Findings.—The anatomical changes found in the adrenal cortex are variable. There may be none (*i.e.* the tissue is overactive but not increased in size); or there may be simple hyperplasia; or an innocent or malignant tumour may be present. (If the tumour cells are non-secreting no endocrine disorder develops.)

The adrenal changes may be primary or secondary. Some cases of Cushing's syndrome are attributed to hypersecretion of pituitary ACTH. An adenoma of the pituitary basophil cells is sometimes present; irradiation of the gland (with X-rays) may lead to regression of the signs of the disease. Changes in the pituitary, consisting of hyalinization of the basophil cells, are said to occur regularly in Cushing's syndrome; but these may be secondary, *i.e.* produced by the elevated blood level of corticoid which is known to act on the pituitary. In some cases of phæochromocytoma (tumours which secrete excess adrenaline and nor-adrenaline) some of the signs of Cushing's syndrome may occur; this is an interesting observation as it is known that adrenaline in adequate concentration stimulates ACTH secretion (p. 950).

1. Cushing's Syndrome ("Corticometabolic Syndrome"). (Hypersecretion of a Cortisone-like Corticoid.)[1]—The main findings are set out below where an attempt is made to interpret them in the light of the known effects of cortisone and related corticoids.

[1] Kepler *et al.*, *Recent Progress Hormone Research*, 1948, 2, 345.

(i) There is wasting and weakness of the skeletal muscles, especially of the limbs. This may be due to the breakdown of muscle protein (to amino-acids) and the low level of serum K^+ (cf. p. 520).

(ii) There is obesity, usually of characteristic distribution. The face is fat (" moon-face "), with narrow eye-slits and a " fish-like mouth " ; there is a " pad of fat " on the back in the cervico-thoracic region ; there is also general trunk adiposity (Fig. 612). The corticoids are known to mobilize and break down reserve fat ; but the clinical findings suggest that in this

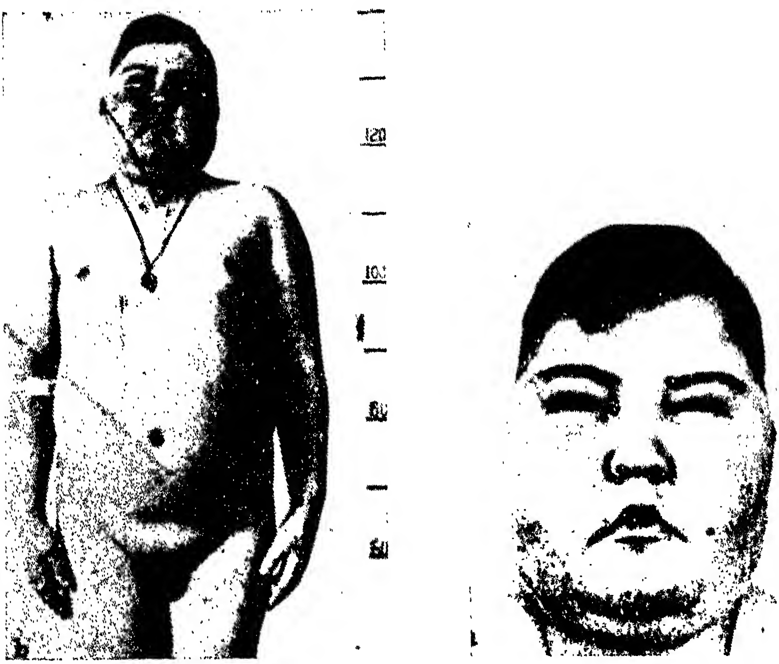


Fig. 612.—Juvenile Case of Cushing's Syndrome. (Melicow and Cahill, *J. clin. Endocrin.*, 1950, *10*, 50.)

Boy aged 11.
Note the moon face, narrow eye-slits and trunk obesity.
The scrotum was small but the testes were well developed.

syndrome a redistribution of reserve fat occurs, *i.e.* mobilization followed by deposition elsewhere.

(iii) The skin is thin and there are purplish striations where the blood vessels show through. The changes are attributed to removal of the protein matrix (cf. (i) *supra*). Acne often develops. There is hirsutism (abnormal hairiness of the face, chest, and abdomen). All these changes are produced clinically by cortisone.

(iv) There is thinning of the bones (osteoporosis) which is due in part to excessive withdrawal of calcium salts and in part perhaps to loss of the protein matrix.

(v) There are signs of diabetes mellitus of varying grades of severity. As explained on p. 961 the cortisone-like corticoids are diabetogenic.

(vi) There are sexual changes: impotence and other signs of hypogonadism in the male and amenorrhœa in the female. These have been attributed to excessive secretion of œstrogen and androgen respectively; but this assumption is unnecessary as cortisone produces similar results.

(vii) The changes in plasma ionic concentration and in the volume and distribution of the extracellular fluids are generally small. Retention of Na^+ and to a less extent of Cl' may occur; alkalosis may result owing to an increase in $NaHCO_3$; serum K^+ is often lowered.

(viii) Hypertension and œdema are common.

(ix) There are changes in the circulating blood cells, *i.e.* a decrease in the eosinophils (p. 952) and in the lymphocytes. Erythræmia may occur (this finding is unexplained).

(x) There is increased excretion in the urine of neutral 17-ketosteroids and corticoids.

In a patient with Cushing's syndrome attributed to a basophil tumour of the pituitary, the following findings were present before, and 9 months after, irradiation of the pituitary.

	Before.	*Nine months after.*
Eosinophils per c.mm.	4	100
Uric Acid/Creatinine in urine	1·10	0·52
Fasting Blood Sugar in mg-% . . .	202	70
Blood Pressure in mm. Hg	230/130	140/95
Urinary Neutral 17-Ketosteroids in mg./24 hr.	230	9·8
Urinary Corticoids in mg./24 hr. . . .	0·65	0·05

Relation of Hypercortism to Œdema and Ascites.—In cardiac œdema (p. 111) and in ascites associated with portal obstruction (p. 822), there is increased reabsorption of Na^+ from the renal tubules into the blood with secondary retention of water. These renal changes have been attributed to excessive corticoid secretion.

2. **Adrenogenital Syndrome.**—As already explained, adrenal cortex extracts contain œstrogens, androgens, progesterone-like substances and cortilactin (a substance acting on the breasts). There is no convincing evidence that the adrenal cortex is *normally* concerned with the regulation of sexual functions. The significance of the presence in the normal cortex of the substances just enumerated, is unknown; it is possible that they may represent intermediaries on the path either of synthesis or of destruction, of the natural cortex hormones.[1] In disease, large amounts of one or more of the sex hormones may be formed in the adrenal cortex with consequent derangement of sexual function. The main clinical syndromes are discussed below.

(1) Precocious Iso-sexual Development in Children.—Thus, little girls have been known to menstruate at the age of two. Little boys show early development of testes, penis, and secondary sexual characters; there

[1] Deoxycortone has progesterone-like qualities; experimentally ovariectomatized monkeys can transform deoxycortone into pregnanediol (the normal end-product of progesterone which is excreted in the urine).

is also precocious growth of the body as a whole resulting in the stocky, precociously virile "infant Hercules" type.[1]

(2) SYNDROMES IN ADULTS.—When the adrenal cortex produces sexual derangements in *adults* there is a tendency to *conversion to the secondary characters of the opposite sex.* Thus adult women with cortical hyperfunction generally develop male secondary sexual characters ; the condition is called adrenal virilism. Adult males, very occasionally, show feminization.

(i) ADRENAL VIRILISM.—The changes produced by adrenal cortex tumours in adult females are illustrated by the following case.[2] The woman was quite normal in appearance till early in 1933 (Fig. 613). Symptoms then developed: she progressively became more masculine in appearance, a beard began to

FIG. 613.—Effects of Tumour of the Adrenal Cortex on Secondary Sexual Characters (Adrenal Virilism). (Hare, Ross, and Crooke, *Lancet*, 1935, ii.)

Left hand : Patient in 1926 at the age of 24. Symptoms began in 1933.
Right hand : Patient in December 1934 at age of 32. Note the remarkable degree of masculinization which has taken place.
A carcinoma of the adrenal cortex was found at operation and death followed its attempted removal.

grow which had to be clipped once weekly, the complexion became florid, the neck thick, and there was a marked double chin. The menstrual periods were first irregular and, later, completely ceased. At operation a *carcinoma* of the adrenal cortex was found. The patient died ; post-mortem the ovaries were small and fibrotic and showed no Graafian follicles ; the uterus was small; the external genitals were normal, but there was some enlargement of the clitoris.

In cases which survive the operation, removal of the adrenal cortex tumour restores the feminine characters, both physical and psychical. The virilism is attributed to excessive secretion of androgens by the adrenal cortex.

(ii) ADRENAL FEMINIZATION.—The commonest change noted in adult males, so afflicted, is enlargement of the breasts ; sometimes considerable glandular development occurs and it may be possible to squeeze out milk

[1] The exploits for which Hercules is justly famous were not exclusively muscular.
[2] Hare, Ross, and Crooke, *Lancet*, 1935, ii, 118.

from the gland. The breast changes are attributed to excessive secretion of œstrogens (which are found in abnormally large amounts in the urine) and probably also of progesterone and cortilactin. Libido is diminished as shown by diminished interest in women, decreased sexual potency and less frequent occurrence of nocturnal erections and of seminal emissions. There may be atrophy of the testes, which is attributed to the raised level of œstrogen in the blood inhibiting the secretion of anterior pituitary gonadotrophins. Successful operation restores normal sexual vigour.

In most cases of the adrenogenital syndrome there is increased urinary excretion of neutral 17–ketosteroids, strikingly so when a secreting carcinoma of the cortex is present (Fig. 609). In cases of feminization the urinary œstrogen content may be high.

(iii) Mixed Syndromes. — In these cases the clinical findings suggest increased secretion of both cortisone-like substances and of sex hormones. Thus the first case of adrenal cortex tumour reported in the literature (by Tilesius in 1803) was a little girl, age 4. She was very obese; her immensely fat cheeks hung over her neck on to the upper part of the chest; the skin was thin, of high colour, with the veins showing through readily; hirsutism was present. But in addition to these "cortico-metabolic" changes there was marked sexual precocity as shown by the development of the external genitals (Fig. 614)

Fig. 614.—Tilesius' Case of Adrenal Cortex Tumour. (Kepler *et al.*, *Recent Progress Hormone Research*, 1948, *2*, 388.) Little girl aged 4. For description see text.

POTASSIUM METABOLISM

Potassium Balance.[1]—The K^+ content of the body depends on the balance of K^+ intake and K^+ loss. As K^+ is an important constituent of all cells, animal and vegetable, any reasonably arranged diet will supply normal K^+ needs. Excretion of K^+ occurs in the urine and some is lost in the fæces.

[1] Elkinton and Tarail, *Amer. J. Med.*, 1950, *9*, 200.

A diet of pure carbohydrate and fat contains no K^+; there is, of course, no K^+ intake during starvation.

The normal K^+ ion concentration in *serum* is about 20 mg-% (5 m.Eq./L) (p. 6); the concentration in the cells is far higher as K^+ is the principal intracellular electrolyte (p. 5). As the surface membrane of most cells is permeable to K^+ ions (p. 7) the great differences which are normally present between the K^+ ion concentration in the intracellular and extracellular fluids are maintained by active cation transfer by the cells, the necessary energy being derived from metabolic processes; when cell metabolism is depressed, *e.g.* in stored red cells, K^+ ions move passively out of the cells into the extracellular fluid with the concentration gradient (p. 8). Most of the intracellular K^+ is combined with organic phosphate esters. In states of *acidosis* these esters undergo phosphorolysis releasing inorganic phosphate which, together with K^+ ions, passes into the plasma (p. 95); as some of the K^+ is excreted in the urine as KH_2PO_4, acidosis leads to depletion of the K^+ reserves of the tissues. In *dehydration* too there is a loss of K^+ from the cells (p. 66).

Regulation of Serum K^+.—Our knowledge of the factors which regulate the serum K^+ level is still imperfect; it is not clear to what extent K^+ can be withdrawn from the plasma and stored in the cells when the serum K^+ rises or can be mobilized from the cells when serum K^+ falls. Serum K^+ depends to an important extent on the degree of K^+ reabsorption which takes place in the renal tubules (p. 28), which in turn depends on the level of adrenocorticoid activity (p. 955).

(i) In Addison's disease serum K^+ often rises; conversely, serum K^+ falls in hypercortism (p. 966) or after injection of DOCA, cortisone or ACTH (pp. 961, 962).

(ii) In familial periodic paralysis (p. 519) a fall of serum K^+ occurs which is roughly proportional to the degree of muscular weakness.

(iii) In renal disease serum K^+ may rise from K^+ retention (owing to decreased glomerular filtration) (p. 4) or fall from excessive excretion in the urine (owing to decreased reabsorption of K^+ in the tubules). Severe K^+ depletion occurs during the phase of recovery from diabetic coma; if the K^+ intake at this time is small, the serum K^+ falls.

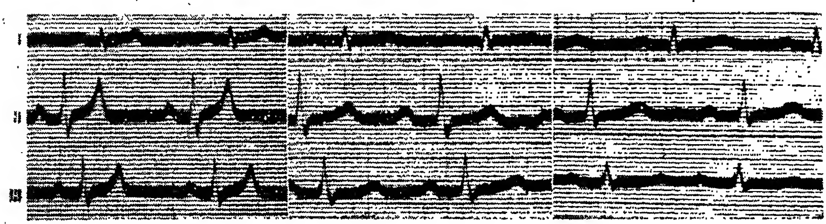
(iv) Serum K^+ may fall owing to K^+ loss in severe vomiting, diarrhœa or steatorrhœa.

Function of Potassium.—The rôle of K^+ is still imperfectly understood. It is related to the excitability of nerve fibre (p. 485) and muscle fibre. It may be involved in transmission processes at the motor end plate (p. 519) and in autonomic ganglia (p. 523). It modifies the action of heart muscle (p. 236). Alterations in the electrocardiogram are a sensitive index of changes in the serum K^+ level (Fig. 615). A fall of serum K^+ causes diminution, depression or occasional inversion of the T wave; the QT interval is prolonged, the ST segment is depressed. An abnormally high serum K^+ level causes the T waves to become exaggerated and sharply peaked; eventually the P waves may disappear.[1] The rôle of intracellular K^+ as a buffer is considered in the case of the red cells on p. 93, and of the general tissue cells on p. 95.

Symptoms of Potassium Lack.—The clinical findings are variable; muscular weakness and the e.c.g. changes mentioned above are the most

[1] Hawkins *et al.*, *Lancet*, 1951, i, 318.

regularly observed. Other symptoms noted in states of K^+ lack and relieved by K^+ therapy are: shallow rapid breathing, abdominal distension and mental changes, *e.g.* lethargy, apathy or delirium.[1] K^+ lack must be treated by giving potassium in amounts sufficient to restore not only the serum K^+ level, but also the intracellular reserves; the total loss of K^+ which occurs in depleted patients may amount to 8–36 g. If the patient can take food by mouth, 3–5 g. of K^+ (as KH_2PO_4 and K_2HPO_4) may be added to each litre



(a) (b) (c)

FIG. 615.—Effect of Changesi n Serum K^+ Level on Electrocardiogram. (Thomson, *Lancet*, 1939, i, 508.

Case of Addison's Disease. Initial blood pressure, 70/45 mm. Hg; blood urea 90 mg–%.

	Serum Levels in mg–%			*Height of T-wave in mm.*		
	K^+	Na^+	Ca^{++}	Lead I	Lead II	Lead III
(*a*) 18th May 1939.	31·6	284	9·9	1·5	7·0	5·0
Put patient on low K^+, high Na^+ diet. Inject extract of whole adrenal cortex ("eucortone").						
(*b*) 10th June 1939.	15·9	309	11·5	1·0	4·0	2·5
(*c*) 23rd June 1939.	8·3	319	9·9	1·0	1·5	0·75

Note in (*a*) the e.c.g. changes resulting from high serum K^+ (hyperkalæmia); note in (*b*) and (*c*) the results of low serum K^+ (hypokalæmia).

of milk or fruit drink; if the flow of urine is adequate such treatment is safe. In urgent cases K^+ may be given parenterally, *e.g.* in Butler's solution (p. 109, footnote) or as Elkinton's solution containing 4·5 g. of K_2HPO_4, 1 g. of KH_2PO_4 and 5·5 g. of NaCl per litre. The rate of infusion should not exceed 20 m.Eq. (0·8 g.) of K^+ per hour, corresponding to 300 c.c. per hour of Elkinton's solution.

Parenteral K^+ therapy may be dangerous if the urinary output is low; the clinical state, the serum K^+ level, and the e.c.g. appearances should be carefully watched for the first signs of potassium poisoning and appropriate measures taken.

THE THYROID[2]

Structure and Chemical Composition.—The normal adult thyroid gland weighs 20–25 g. It consists of many spherical or oval vesicles which are lined by a cubical epithelium (Fig. 616, B). The cavity of the vesicles is filled with a viscid homogeneous material known as the *colloid* of the gland; the colloid consists chiefly of an iodine-containing protein called *thyroglobulin.*

[1] Eliel *et al., New Engl. J. Med.*, 1950, *243*, 471, 518.

[2] Salter, *The Hormones*, N.Y., 1950, *2*, 181, 301. Pitt-Rivers, *Physiol. Rev.*, 1950, *30*, 194. Symposium on "Thyroid function as disclosed by newer methods of study," *Ann. N.Y. Acad. Sci.*, 1949, *50*, 279–508. Pochin, *Lancet*, 1950, ii, 41, 84. Means, *Thyroid and its Diseases*, Phila., 1937. Harington, *Thyroid Gland*, London, 1933.

The thyroid gland contains on an average 2 mg. of iodine per g. dry weight (1 g. dry=4–5 g. wet weight); the weight of iodine found in the gland represents 20% of the total body iodine. The iodine of the gland is normally found in three main forms in the proportions indicated: inorganic *iodide* (10%); (there is no free iodine in the gland); *diiodotyrosine* (60%); *thyroxine* (30%). Thyroxine is the active principle of the gland and is secreted as such into the blood; diiodotyrosine is its principal precursor in the gland.

Thyroglobulin contains both thyroxine and diiodotyrosine bound to the rest of the protein molecule by the usual peptide (—CO.NH—) linkage. Thyroglobulin is thus a storage form of thyroxine. The thyroid differs from the other ductless glands in storing its active principle in the cavity of a vesicle rather than in the secreting cells themselves; when required, the stored (combined) thyroxine is digested off from the thyroglobulin molecule by proteases and is secreted, in the free form, into the blood.[1]

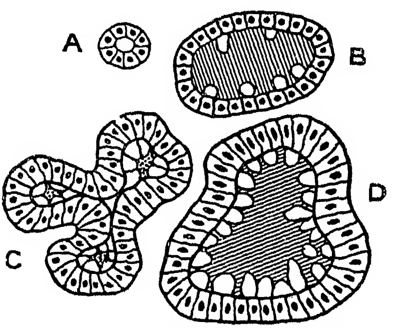


FIG. 616.—Histological Appearance of Alveoli of Thyroid under various Conditions. (After Means, *Thyroid and its Diseases*.)

A. Fœtal type, containing no colloid.
B. Normal alveolus lined by a cubical epithelium; it is filled with colloid which contains few vacuoles.
C. Hyperplastic alveolus (*e.g.* of exophthalmic goitre); the lining cells are columnar and increased in number; the colloid has been discharged. The wall of the alveolus is infolded.
D. Hyperplastic alveolus (*e.g.* as in C) after treatment with iodine. The alveolus has become distended with retained colloid; numerous vacuoles are present in the colloid.

The size of the thyroid, its histological appearance, the amount of colloid, and its chemical composition vary widely under different conditions. The points to be noted are: the size and shape of the vesicles; the height of the epithelium; the total iodine content of the colloid; and the relative amounts of the three main iodine-containing constituents.

(i) When the gland is *stimulated* it becomes more vascular, the lining cells become more numerous (*hyperplasia*) and the circumference of the vesicles in cross-section is increased; the epithelial cells become taller. It is customary to call such an overactive gland a *hyperplastic* gland.

(ii) When rapid *secretion* is taking place, the colloid is digested away and becomes first faintly staining and vacuolated and later reduced in amount leading to an *infolding* of the proliferated epithelium (Fig. 616, C).

(iii) The administration of *iodine* often produces characteristic changes in a hyperplastic gland (whatever the cause): the amount of colloid is greatly increased (*i.e.* there is increased "storage") and the epithelium regresses to the normal state (Fig. 616, D).

(iv) In *atrophy* of the gland the epithelium becomes thinner and the vesicles smaller and fewer in number.

The normal daily iodine intake in the food is of the order of 20–200 μg.[2]; it is, of course, higher if iodine is artificially added to the diet.

[1] But diiodotyrosine is also present in the blood.
[2] 1 μg.=1 microgram=0·001 milligram.

Thyroxine Formation and Secretion.—The main processes which take place in the thyroid gland are:

(i) Selective absorption of inorganic iodide from the plasma.

(ii) Oxidation of the iodide to iodine.

(iii) Immediate binding of the newly formed iodine with tyrosine to form diiodotyrosine.

(iv) Transformation of diiodotyrosine into thyroxine.

(v) Secretion of thyroxine (*a*) into the blood, or (*b*) into the lumen of the vesicles, to be stored in the colloid.

All these processes doubtless depend on the activity of various enzyme systems, none of which have as yet been identified. The secretion of thyroxine

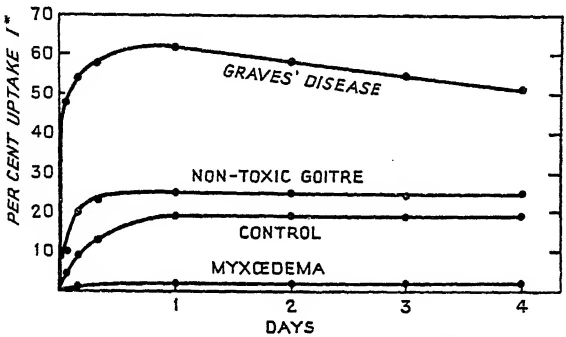


FIG. 617.—Uptake of Radio-active Iodide ($^{131}I^*$) by Hypothyroid, Normal, and Hyperthyroid Glands. (Hamilton, *Radiology*, 1942, *39*, 555.)

A tracer dose of radio-active iodide was given to human beings. Note the negligible iodide uptake by the hypothyroid gland and the excessive uptake by the hyperthyroid glands.
Radio-iodide concentration was determined by means of a Geiger counter on the neck over the gland.

is regulated (*a*) mainly by the thyrotrophic hormone of the anterior pituitary; (*b*) possibly by the level of blood thyroxine and blood iodide; (*c*) to a minor degree by the sympathetic nerve supply (p. 982). Most of the stages referred to above need more detailed consideration.

(1) SELECTIVE ABSORPTION OF IODIDE.—This process is well demonstrated in experiments in which minute (*tracer*) doses of radio-active iodine (usually $^{131}I^*$) are given in the form of iodide[1]; the doses administered are comparable to the amounts which are normally ingested in the diet. The distribution and fate of the labelled iodide can be easily followed. It is rapidly absorbed from the intestine and passes into the plasma; the plasma $^{131}I^*$ consequently rises initially. The $^{131}I^*$ is selectively taken up by the thyroid gland; *e.g.* in man 20% of the administered $^{131}I^*$ may be present in the gland within 6 hours (Fig. 617, Control).

A normal thyroid "collects" up to 80 times the amount of $^{131}I^*$ that would be expected were the iodide diffused uniformly throughout the body. A hyperplastic gland may collect up to 300 times the amount which would

[1] $^{131}I^*$ has a half-life of 8 days.

be expected from simple diffusion. These data prove that the thyroid cells selectively absorb inorganic iodide and retain it in spite of a steep diffusion gradient that tends to drive the iodide back into the plasma. It is noteworthy, however, that there is an absolute maximum iodide content of the gland no matter how great are the iodide supplies, *i.e.* there is a thyroid "iodide absorption maximum" which cannot be exceeded.

In man the radio-active I content of the thyroid gland can be determined by means of a Geiger counter placed on the neck, over the gland. On administering tracer doses of $^{131}I^*$, the iodide uptake of the gland is found to be negligibly small in myxœdema but markedly enhanced in hyperthyroidism (Fig. 617).

As is shown below the $^{131}I^*$ is later incorporated first into diiodotyrosine and then into thyroxine.

The thyroid gland is the only organ in the body which can selectively absorb inorganic iodide from the plasma. This process is paralysed by thiocyanates (cf. phloridzin which paralyses the reabsorption of glucose in the renal tubule). After administration of KCNS the inorganic iodide content of the gland is low and is only increased when the blood iodide is markedly raised by giving large amounts of KI.

Before the thyroid iodide can become organically bound it must be first converted by oxidative enzyme systems into free *iodine.*

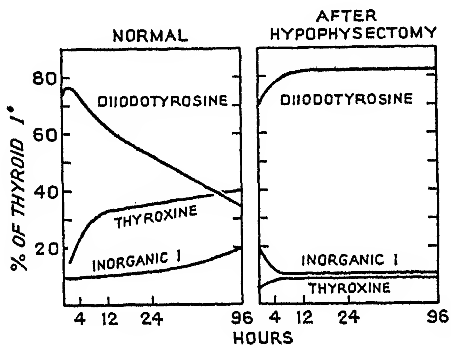


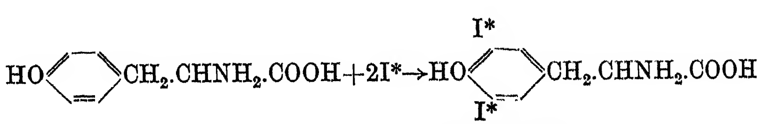
FIG. 618.—Iodine Metabolism in Normal Thyroid and after Hypophysectomy. (Pearlman *et al.*, *Endocrin*, 1942, *30*, 497.)

I* = radio-active iodine ($^{131}I^*$). Experiments in rats following administration of $^{131}I^*$ (as iodide).

Left-hand Figure.—*Normal rat*: distribution of $^{131}I^*$ among inorganic iodide, diiodotyrosine and thyroxine fractions in thyroid gland. Note decline of diiodotyrosine and progressive increase in thyroxine concentrations as the former is transformed into the latter.

Right-hand Figure.—*Hypophysectomized rat*: Note initial increase in diiodotyrosine, but there is no subsequent decline as *no thyroxine is formed.*

(2) FORMATION OF DIIODOTYROSINE AND THYROXINE.—Tracer experiments show that the iodide taken up by the gland, immediately after its oxidation to iodine, is first bound with tyrosine to form diiodotyrosine.

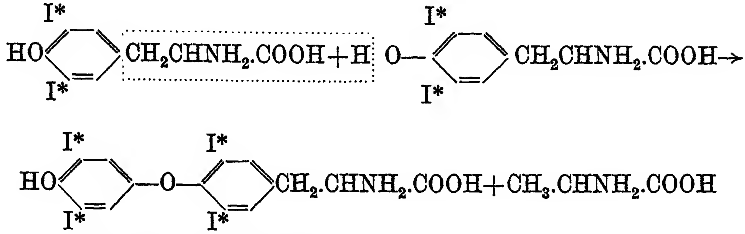


Tyrosine Diiodotyrosine

$I^* = {}^{131}I^*$.

The diiodotyrosine concentration in the gland then progressively falls and the thyroxine concentration gradually rises as the former is transformed into the latter compound (Fig. 618). The formation of thyroxine may

result from the condensation of two molecules of diiodotyrosine with the removal of alanine thus :



Thyroxine (para-hydroxy-phenyl ether of tyrosine plus 4 I atoms) Alanine

It is presumed that diiodotyrosine and thyroxine formation occurs in the epithelial cells of the gland.

(3) Rôle of the Anterior Pituitary.—After hypophysectomy (in rats) thyroxine metabolism is disordered. Inorganic iodide is taken up by the gland from the plasma more slowly and to a smaller extent ; diiodotyrosine is, however, formed normally. The essential disturbance is the almost complete arrest of the normal transformation of diiodotyrosine into thyroxine. Only minimal amounts of thyroxine are formed and secreted into the blood.[1] The anterior pituitary (through its thyrotrophic hormone) has two main actions : (i) it activates the enzyme systems responsible for the conversion of diiodotyrosine into thyroxine ; (ii) it stimulates the secretion of thyroxine (Fig. 618).

(4) Action of Thio-Compounds.—Many organic thio-compounds containing the chemical group —N.CS.N— (*e.g.* thiouracil) interfere with thyroxine formation in the gland. The absorption of inorganic iodide by the gland takes place normally but it is not converted into diiodotyrosine or thyroxine (cf. p. 996). These thio-compounds thus block the binding of iodine by tyrosine.[2]

(5) Action of Inorganic Iodide Level.—Using thyroid slices, it is found that if the level of inorganic iodide in the medium is raised to 10–15 μg./100 c.c. (normal serum level is 3 μg.), synthesis of both diiodotyrosine and thyroxine is depressed.

(6) Formation of Thyroxine outside the Thyroid Gland.—Thyroxine can be formed in small amounts in the general tissues of the body. The evidence is as follows :

(i) Radio-active iodide is administered to completely thyroidectomized animals ; it is subsequently found in the bound form in diiodotyrosine and in thyroxine, both in the tissues and in the plasma. The basal metabolic rate is raised because of the specific activity of the newly formed thyroxine. It is claimed that thyroxine synthesis can even occur in the hypophysectomized, thyroidless animal.

Although these observations on the thyroidless animal are true, the thyroid still remains the only significant site of thyroxine formation, probably

[1] As explained on p. 931, in clinical severe hypopituitarism, some thyroid secretion continues.

[2] A sharp distinction should be drawn between the *organic* thio-compounds like thiouracil and *inorganic* thio-compounds like KCNS, which is discussed above (p. 974).

for two reasons: (*a*) it has a *uniquely high iodide content* which favours thyroxine formation; (*b*) it may also contain an optimal concentration of the various *enzymes* which catalyze thyroxine formation.

(ii) Thyroxine can be synthesized *in vitro* with surprising ease: (*a*) If diiodotyrosine-peptone is incubated with pepsin, it is built up into a protein containing thyroxine; (*b*) If serum proteins or casein are incubated with iodine (in unphysiologically high concentrations, in an alkaline medium and at a high temperature), thyroxine appears (presumably derived from the tyrosine content of the protein); the ultimate product may be more active than thyroglobulin. It must be remembered, however, that the thyroid gland forms thyroxine under *physiological* conditions of iodine concentration, *p*H and temperature.

When dried thyroid is administered by mouth its physiological activity is greater than that produced by administering its thyroxine content alone. It seems probable that the diiodotyrosine (in protein combination) of the dried gland material (representing about 60% of the total iodine content) is transformed somewhere in the body into thyroxine. (Pure (*i.e.* uncombined) diiodotyrosine given by mouth is inert.)

Cyclic Activity of the Thyroid.—The different thyroid vesicles are not all in the same phase of activity at any one moment. They may be carrying out one or two of the following three processes: (i) forming and secreting thyroxine into the blood; (ii) storing thyroxine in the colloid; (iii) mobilizing thyroxine from the colloid store for secretion into the blood.

Blood Iodine.—Iodine is found in the blood mainly in two forms:

(1) As *inorganic iodide* (so-called *filtrable iodide*). On a normal iodine intake the inorganic iodide in plasma and corpuscles is less than 1 μg–% (mean, 0·5 μg.). Its concentration bears no relationship to the level of activity of the gland but it immediately reflects changes in iodine intake. Thus the ingestion of 15 grains (=1 g.) of KI raised the serum inorganic iodide, 4 hours later, to 2300 μg–% (=2·3 mg–%); this is the level that one expects to find if the ingested iodide is uniformly diffused throughout the body water (=50L). After 50 hours the serum inorganic iodide had fallen to 46 μg–% owing to steady excretion of iodide in the urine.

(2) As *hormone iodine* (*i.e.* thyroxine and also diiodotyrosine) present in the plasma only, and loosely bound with the serum albumin; this iodine is precipitated out with the serum proteins. It is, therefore, also called *precipitable* or *protein-bound iodine*. The normal range of hormone iodine is 4–8 μg–%. The serum hormone iodine level is a reliable index of thyroid secretory activity.

(i) In myxœdema its level is 0·2–2·5 μg–% (mean 1·3) and is restored to normal by thyroid treatment.

(ii) After total thyroidectomy in man the level falls to about zero, indicating that hormone formation outside the thyroid gland in man is negligible.

(iii) In hyperthyroidism the average serum hormone iodine is 14 μg–% and values as high as 30 μg–% have been recorded (Fig. 628).

(iv) Partial thyroidectomy reduces these high values to normal (or subnormal) depending on the amount of gland excised. Iodine therapy which reduces thyroid secretion in Graves' disease, likewise lowers the hormone iodine level (though it raises the filtrable, *i.e.* the inorganic iodide level).

Actions of Thyroxine.—(1) General Metabolism.—The outstanding

action of thyroxine is to *stimulate metabolism* in the tissues generally, increasing oxygen consumption, presumably by catalyzing the enzyme systems which are responsible for oxidation processes. The exact way in which thyroxine produces its effect is unknown except for the fact that it acts *directly* on the tissues. Thyroxine also promotes the *growth and development of the organs generally*; this latter action is best demonstrated in the young thyroidectomized animal (cf. cretinism, p. 986).

(i) *Action in Myxœdema.*—The effect on metabolism is most simply studied in the hypothyroid patient. In one case of myxœdema, the basal metabolic rate was 30% below normal (*i.e.* the B.M.R. was −30); 16 mg. of thyroxine were given by mouth in one dose. The B.M.R. (*i.e.* oxygen consumption, energy output) rose gradually to normal and then slowly fell to its original level which was reached at the end of six weeks. During this period the extra energy output due to the administered thyroxine was 16,000 Cal.; each mg. of thyroxine thus increased metabolism by an amount equivalent to the oxidation of 250 g. of glucose (*i.e.* 1000 Cal.). About 1 mg. of thyroxine daily (=3 grains of dry thyroid), restores the B.M.R. in a severe case of myxœdema to normal; this result is an indication of the normal rate of secretion of the hormone.

Thyroxine has a long period of latency; even after the intravenous injection of pure thyroxine in man 48 hours may elapse before there is a detectable effect. If one large dose of thyroxine is given there is considerable loss by excretion in the bile, and results are obtained more economically with small repeated doses.

It is customary clinically to treat hypothyroid patients by means of dried thyroid given by mouth. This preparation is more active than would be expected from its thyroxine content because (as mentioned on p. 976) some of its contained diiodotyrosine is also converted in the body into thyroxine. The digestive juices do not inactivate the thyroid active principles (all the other body hormones, however, are less effective by mouth than by subcutaneous or other parenteral routes).

Recently synthetic L-thyroxine sodium has been prepared sufficiently cheaply to enable it to compete in price with thyroid extract. It has been used successfully in a series of cases of myxœdema (the daily oral maintenance dose was 0·15–0·3 mg.).[1] This preparation has the advantage over thyroid extract that it does not need biological or chemical standardization.

(ii) *Action in Normal Subjects.*—If thyroxine or dried thyroid is given to *normal* subjects in doses which cure myxœdema, the B.M.R. rises much less than would be expected from the results observed in myxœdema; in some cases the rise of B.M.R. produced in the normal subject may be negligible and there is no increase in the serum hormone iodine level. The explanation of the discrepancy may be: (*a*) that there is a corresponding compensatory decrease of thyroxine secretion by the normal gland; (*b*) that the thyroid takes up the excess thyroxine from the blood and destroys or stores it.

If dried thyroid is given in *very large* doses, *e.g.* 10 grains daily (the therapeutic dose in severe myxœdema is 3 grains daily), the serum hormone iodine level rises and many of the characteristic signs of hyperthyroidism develop, *e.g.* rise of B.M.R., loss of weight, fall of serum cholesterol level, and rapid heart rate (Fig. 619). Exophthalmos is, however, never produced;

[1] Hart and McLagan, *Brit. med. J.*, 1950, i, 512.

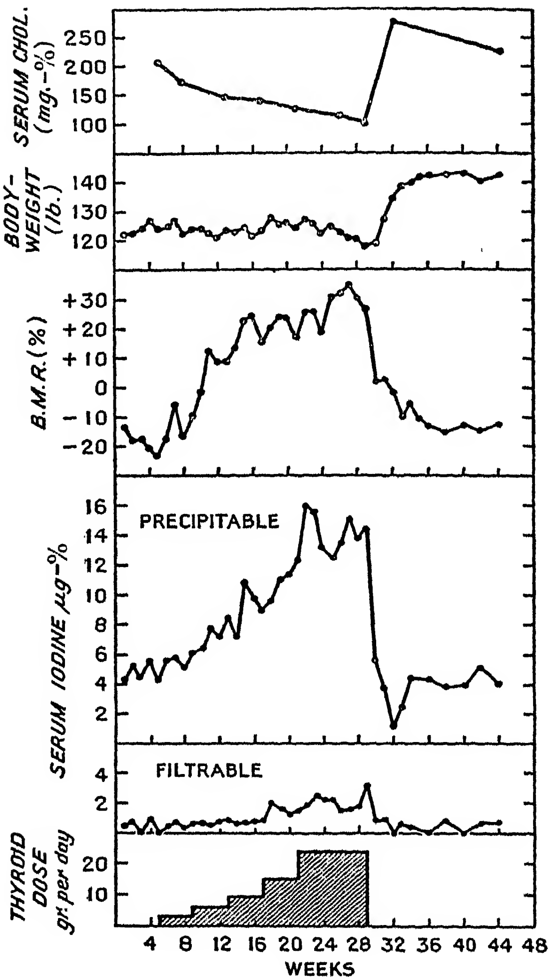


FIG. 619.—Effects of Administration of Excessive Doses of Thyroid in Normal Person. (Riggs *et al.*, *J. clin. Investig.*, 1945, *24*, 723.)

Records from above downwards are:
Serum cholesterol in mg/100 cc.
Body weight in lbs.
BMR as per cent. above or below normal.
Serum precipitable (hormone) iodine in µg-%.
Serum filtrable iodine in µg-%.
Dose of thyroid administered in grains per day.
The subject was a woman with schizophrenia but otherwise "euthyroid." Note, however, that during the control period the B.M.R. fluctuated between −10 and −25.
The dose of thyroid was increased at intervals of 4 weeks from 3 to 6, 10, 15, 20, and finally 25 grains per day. The highest dose was maintained for 8 weeks and then abruptly discontinued.
Note the progressive return of the various values to normal.

this point is important in relation to the causation of exophthalmos in Graves' disease.

(2) Special Metabolism.—Thyroxine promotes the absorption of glucose from the small intestine (p. 836); in the liver it stimulates the conversion of glycogen into glucose and promotes glucose formation from non-carbohydrate sources. For all these reasons thyroxine produces a transient rise of blood sugar and glycosuria (*thyroid diabetes*). Large doses of thyroxine may produce lasting damage to the islets of Langerhans leading to persistent hyperglycæmia and glycosuria (*metathyroid diabetes* (p. 918)).

(3) Heart. — Thyroxine accelerates the normal and the denervated heart; it thus has a direct peripheral action on the heart. Under natural conditions part of the accelerator action of thyroxine is secondary to the raised metabolism.

(4) Bone.—Thyroxine modifies the metabolism of *calcium*. It causes removal of calcium (together with phosphate) from the bones, leading to their rarefaction (osteoporosis). Unlike parathormone it does not raise the serum calcium, and increased Ca loss occurs in the fæces as well as in the urine (cf. p. 1003). The effect is not due to the increase in metabolic rate, because no such changes take place when, for example, the metabolism is increased by fever. Similar changes in the bones are observed in Graves' disease (p. 991). The mode of action of thyroxine on calcium metabolism is unknown; it is, however, antagonized by administration of vitamin-*D*.

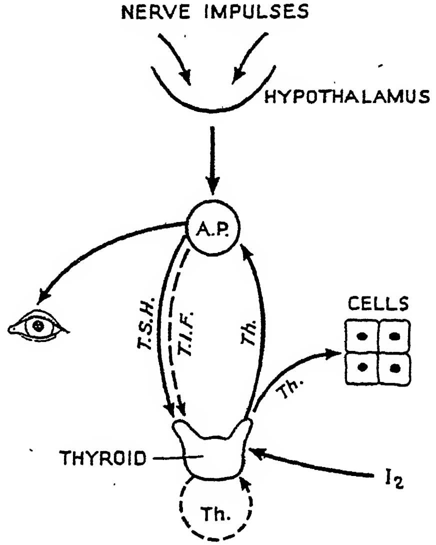


Fig. 620.—Regulation of Thyroid Secretion. (After Means and Soskin.)

A.P., anterior pituitary.
T.S.H., thyroid stimulating hormone (thyrotrophin).
T.I.F., hypothetical thyroid inhibiting hormone.
Th., blood thyroxine.
I_2, blood iodine level.

(5) Kidney.—The action of thyroxine on the flow of urine, especially in diabetes insipidus, is discussed on p. 51. It probably also regulates the distribution of fluid in the body; thus in myxœdema there is excessive fluid retention in the extracellular spaces (p. 983).

(6) Breast.—Thyroid may promote the secretion of milk (p. 1094).

Regulation of Thyroid Secretion. (Fig. 620).—The main regulator of thyroid secretion is the *anterior pituitary* via its *thyrotrophic hormone* (*TSH*, *thyroid-stimulating hormone, thyrotrophin*). The release of thyrotrophin is determined mainly by the level of *blood thyroxine* acting directly on the pituitary, and possibly also by various mechanisms which act reflexly via the *hypothalamus*. Other factors which influenced thyroid secretion are: (i) the level of *blood iodide*; (ii) variations in *external temperature*; (iii) possibly

the *sympathetic* nerve supply to the gland. Some workers claim that the anterior pituitary may sometimes secrete a thyroid-*inhibiting* hormone.

(1) **Rôle of Anterior Pituitary. Thyrotrophin**—(i) RESULTS OF HYPOPITUITARISM—Hypophysectomy leads to thyroid atrophy; the thyroid is restored to normal by pituitary grafts. Similarly, considerable thyroid atrophy occurs in the classical clinical condition of general anterior pituitary failure known as Simmonds' disease; in this condition, however, the gland continues to secrete at a low steady rate. The B.M.R. falls to −25 compared with −40 after complete thyroidectomy (p. 931).

(ii) THYROTROPHIN has been isolated from the anterior pituitary in a highly active form; it is a glycoprotein. It is thought to be secreted by the basophil cells of the gland; in conditions associated with excess thyrotrophin secretion the basophils increase in number and the eosinophils almost disappear. Injection of thyrotrophin produces characteristic thyroid hyperplasia (*i.e.* increased vascularity, increased height of the epithelium, cellular proliferation, decreased colloid content of the vesicles), increased secretion of thyroxine, and increased metabolic rate. The histological appearance resembles that of the gland in Graves' disease.

(iii) FATE OF THYROTROPHIN.—The hormone combines specifically with thyroid tissue and is secondarily inactivated by it. It is more rapidly removed from the blood by a hyperplastic thyroid; for this reason the concentration of thyrotrophin in the blood (or its rate of excretion in the urine) is not necessarily a reliable guide to its rate of formation. The only other tissues which can inactivate thyrotrophin are the thymus, the lymph glands, and possibly the extraocular structures.

(iv) CONTROL OF THYROTROPHIN SECRETION.—The control is partly nervous via the hypothalamus and partly determined directly by the level of blood thyroxine.

(*a*) *Nervous Control.*—There is good evidence that stimulation of the hypothalamus causes the release of several of the anterior pituitary hormones (p. 931). It is possible that the hypothalamus also causes secretion of thyrotrophin. It is believed by some that emotional tension causing hypothalamic overactivity leads to oversecretion of thyrotrophin and the development of Graves' disease.

(*b*) *Rôle of Blood Thyroxine.*—The secretion of pituitary thyrotrophin is regulated by the level of blood thyroxine; thus it is stimulated by a low blood thyroxine and inhibited by a raised blood thyroxine level. Normally the activity of the pituitary and thyroid are finely integrated to maintain an appropriate blood thyroxine level with its resulting effects on the activities of the general tissues of the body.

In any condition in which the blood thyroxine level is kept abnormally low because of reduced thyroxine secretion or because of increased thyroxine utilization, there is *increased secretion of pituitary thyrotrophin* and consequently *thyroid hyperplasia* (*if thyroid tissue is present and is responsive*). Such induced thyroid hyperplasia always regresses after hypophysectomy.

Thyroxine decreases the oxygen consumption of thyroid slices though it increases the oxygen consumption of all the other tissues; *excess blood thyroxine may thus also directly depress thyroid secretion.* A fall of blood thyroxine produces the reverse effects.

(v) THYROID HYPERPLASIA INDUCED VIA THYROTROPHIN BY LOW BLOOD

THYROXINE LEVEL.—Hyperplasia caused in this way may occur in the following conditions:

(*a*) *Iodine insufficiency* in the diet (p. 986).

(*b*) *Poisoning with organic thio-compounds* (*e.g.* thiouracil) and similarly acting substances, which inhibit the iodination of tyrosine in the thyroid and so prevent thyroxine formation. The thyroid undergoes progressive enlargement, *e.g.* from 5 to 25 mg. per 100 g. of body weight and shows characteristic hyperplasia. The thyroid decreases in size after hypophysectomy (p. 989).

(*c*) *Poisoning with KCNS* (thiocyanate) which interferes with iodine uptake by the gland and therefore with thyroxine formation. In a clinical case of this condition the thyroid was enlarged and hyperplastic owing to the compensatory pituitary overactivity; there were (as would be expected) clinical signs of *hypothyroidism* (*e.g.* B.M.R. of —18); the serum hormone iodine level was depressed (to 2·1 μg-%). On administering thyroxine the blood thyroxine (and B.M.R.) rose and the thyroid became smaller.

(*d*) *Partial thyroidectomy* results in decreased thyroxine formation with consequently increased secretion of thyrotrophic factor; the remaining portion of the thyroid enlarges and shows hyperplasia, and its colloid and iodine content are reduced as thyroxine secretion is stimulated. After complete removal of the thyroid (and in severe myxœdema) the blood and urine levels of thyrotrophin are increased (but of course there is no thyroid tissue for thyrotrophin to act upon).

(*e*) The thyroid commonly enlarges during adolescence, in the latter part of the menstrual cycle and during pregnancy. The enlargement may be secondary to increased thyroxine utilization by the reproductive organs.

(vi) ACTION OF THYROTROPHIN ON EXTRAOCULAR TISSUES.—In many species anterior pituitary extracts produce exophthalmos. In experiments on fish the exophthalmos is due to a great (fivefold) increase in the amount of fluid in the retrobulbar space. As the ocular changes also develop in the thyroidless animal they are not mediated by thyroxine. It is suggested that the extracts have a direct action on the orbital tissues causing (in some unexplained way) more extracellular fluid to be formed than can be removed by absorption; there is swelling not only of the retrobulbar tissue but also of the *extraocular muscles* and the conjunctiva.

It is generally supposed that the pituitary hormone (thyrotrophin) which stimulates the thyroid is identical with the substance which produces this experimental exophthalmos.

There is great uncertainty about the action of thyrotrophin on the eye structures in *man*. The possible relation of thyrotrophin to the eye changes in Graves' disease is considered on p. 993.

(vii) It has been suggested that the anterior pituitary may under some circumstances secrete a thyroid-*inhibiting* factor. Thyroid was administered to rats; injection of extracts of their anterior pituitaries into guinea-pigs reduced the size of the thyroid, the epithelium became lower, the vesicles were distended with colloid and the B.M.R. was reduced. (In control experiments, extracts of normal rat anterior pituitary produced characteristic hyperplasia.)

(2) **Rôle of Level of Blood Inorganic Iodide.**—The level of blood

inorganic iodide depends on the *iodine intake* and its rate of utilization and excretion.

(i) An adequate iodine intake is necessary for normal thyroxine formation.

(ii) Iodine deficiency depresses thyroxine formation and leads to thyroid hyperplasia combined with hypothyroidism (p. 987).

(iii) Administration of excess iodine has no effect on the *normal* gland.

(iv) In most states of thyroid *hyperplasia*, however, iodine treatment leads to characteristic changes: there is increased accumulation of a colloid which is rich in iodine; the hyperplasia regresses, *e.g.* the height of the epithelium is reduced, and less thyroxine is discharged into the circulation. Thus a high blood-iodide level favours colloid *storage* rather than hormone discharge; in fact it impedes hormone secretion. In Graves' disease, iodine medication reduces thyroxine secretion and so improves the clinical state; the vesicles which were previously empty, become filled with colloid. Similarly the initial effect of administering small doses of iodine in cases of iodine-deficiency goitre is to cause colloid accumulation and consequently, further enlargement of the goitre (p. 987).

The way in which a raised blood inorganic iodide level produces its effects in hyperplasic states is unknown.

(3) **Rôle of Sympathetic Nerves.**—The influence of the sympathetic on thyroid secretion is probably unimportant. No changes in the structure of the gland or in the B.M.R. follow denervation of the thyroid. It is claimed, however, that sympathetic nerve stimulation decreases the thyroid iodine content and increases the blood iodine concentration.

(4) **Rôle of Body Temperature.**—The rate of thyroid secretion is affected by environmental temperature; a fall of external temperature increases the secretion. Thus in the rat, if the rate of thyroid secretion at 35° C. is taken as unity, the rate increases to 3·0 when the external temperature is lowered to 25° C. and to 5·5 when it is lowered to 1° C. The extra thyroxine which is released stimulates the metabolic rate, increases heat production and so helps to maintain body temperature. Characteristic signs of thyroid hyperplasia develop in cold environments.

Valuable information about the physiology of the thyroid gland in man is obtained from a study of clinical cases of thyroid deficiency (myxœdema in adults and cretinism in children) and of thyroid excess (hyperthyroidism, Graves' disease).

Thyroid Deficiency in Man.—1. Myxœdema.—In *man*, total extirpation of the thyroid gland in adult subjects (*e.g.* for malignant growth) produces *post-operative myxœdema*; a similar result may follow excessive removal of thyroid tissue in patients with Graves' disease (cf. Fig. 623). The symptoms set in within a few weeks or months after the operation. An identical clinical condition of *myxœdema* occurs spontaneously in adults—especially in women—from atrophy of the thyroid gland (Fig. 621). Minor forms of this disease are common; the gland is usually small, the vesicles are few and contain little colloid.

The chief symptoms can be readily understood if it is remembered that owing to lack of thyroxine many of the activities of the body are depressed.

(i) There is a diminution of mental activity, loss of memory, and slowness of thought, speech, and movement; there is considerable muscular weakness.

(ii) The skin becomes dry and coarse; the cheeks are parchment-like in

appearance, and there is a slight malar flush. The hair is coarse and tends to fall out of the head and the outer thirds of the eyebrows. The subcutaneous tissues are infiltrated with excess fluid containing a muco-protein giving a puffy appearance to the body. The capillaries in myxœdema are said to be excessively permeable to proteins. The œdema is characteristically firm and "solid."

(iii) The sexual functions are disturbed; amenorrhœa commonly develops in women.[1]

A B

Fig. 621.—Treatment of Myxœdema with Thyroid. (Harington, *The Thyroid Gland*, Oxford University Press, 1933.)

A. Patient, aged 65, bedridden and imbecile from myxœdema of twenty years' standing.
B. Same patient after fifteen months' treatment with thyroid. The regrowth of hair is a striking feature. The health was normal, and she was happy and mentally alert. She died at the age of 94 after twenty-nine years of vigorous life under treatment. (Raven's case.

(iv) The basal metabolic rate (B.M.R.), *i.e.* the energy output of the individual at complete rest 12–18 hours after a meal (p. 377), is diminished to 30 or even 45% below normal (B.M.R. is —30 or —45); in other words, the rate of oxidation in the tissues is greatly diminished. The pulmonary ventilation, the oxygen consumption, the CO_2 output, the nitrogenous excretion in the urine are all diminished, because the general metabolic activities are depressed. As the tissues need less oxygen, the output of the heart is reduced, *e.g.* to 2·5 litres per minute. This is due in part to a decrease in the

[1] In monkeys, thyroidectomy results in amenorrhœa; the menstrual periods are restored by thyroid treatment.

heart rate and in part to a diminution of the output per beat. As heat production is decreased the body temperature may be subnormal (p. 474); the patients show increased sensitivity to cold and react less well than normal subjects to a cold environment.

(v) Body weight increases because of decreased metabolism and the extensive œdema in the skin and internal organs.

(vi) Certain blood changes occur regularly in myxœdema (Fig. 623).

(*a*) The serum *cholesterol* is increased from the upper limit of the normal of

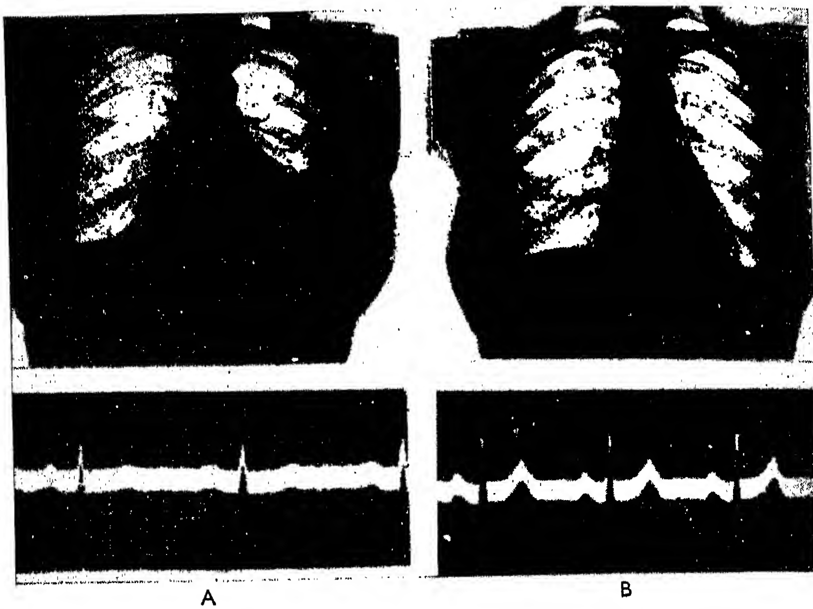


FIG. 622.—Heart Changes in Human Myxœdema. Effects of Thyroid Therapy. (Master and Stricker, *Ann. int. Med.*, 1941, *15*, 124.)

A. Teleroentgenogram and electrocardiogram (Lead II) in case of post-operative myxœdema, BMR, −20: heart rate 70; BP, 100/70: some precordial pain and shortness of breath; heart "water-bottle-shaped" (enlargement of transverse diameter): low voltage QRST waves.

B. After two months of thyroid treatment: the heart is smaller and contracts vigorously; increased QRST voltages; there is slight left axis deviation. The heart rate is increased.

300 mg–% to, for example, 400 or even 700 mg–%. The blood fatty acids and phosphatides follow the serum cholesterol level, *i.e.* there is a general lipæmia. The changes in serum cholesterol are well correlated with the clinical state and with the level of the B.M.R. It is believed that a serum cholesterol level below 275 mg–% excludes a diagnosis of myxœdema.

(*b*) The plasma *protein* level tends to be raised.

(*c*) The serum *hormone iodine* (precipitable iodine) is diminished (p. 976).

(*d*) *Anæmia* commonly occurs; it may be due to iron deficiency or other causes, but sometimes it is the specific effect of lack of thyroid secretion leading to depressed activity of the bone marrow. In such cases the anæmia is relieved only by thyroid treatment (p. 194).

(vii) The changes in the heart are typical (*myxœdema heart*). The heart is enlarged in its transverse diameter, is flabby, and contracts sluggishly. The ventricular waves (QRST) in the electrocardiogram are of low voltage. The cardiac output (as stated) is decreased. It is suggested that the heart's function is disturbed by œdema of the muscle fibres (Fig. 622).

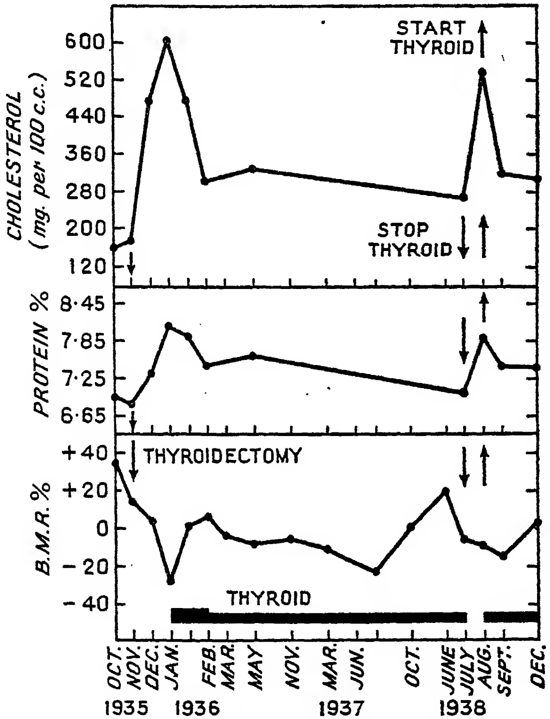


Fig. 623.—Changes in Blood and Metabolism in Post-operative Myxœdema. Results of Thyroid Therapy. (After Gildea *et al.*, *J. clin. Investig.*, 1939, *18*, 747.)

Ordinates from above downwards:
Serum cholesterol in mg. per 100 cc.
Plasma protein in g–%.
Basal metabolic rate as per cent. above or below normal. 0 = normal.

The patient was first seen on October 29, 1935, with signs of *hyperthyroidism*: serum cholesterol 160 mg–%; B.M.R., +40.

At first arrow (November 17, 1935) *partial thyroidectomy* was performed which was followed by post-operative *myxœdema*.

Note: (i) Rise in serum cholesterol (to 600 mg–%.). (ii) Rise in plasma proteins (to 8·2%). (iii) Fall in B.M.R. (to −30). (iv) Body weight rose and œdema developed.

Thyroid treatment was begun on January 9, 1936, 1½ grains daily to February 13, and 1 grain daily subsequently.

Note: (i) Fall in serum cholesterol to about 300 mg–%. (ii) Fall in plasma protein to about 7%. (iii) Irregularly maintained increase in B.M.R.

On July 14, 1938, thyroid treatment was stopped.

Note: The characteristic blood and B.M.R. changes of myxœdema returned.

On August 25, 1938, thyroid treatment resumed.

Note: Signs of myxœdema cleared up.

All the symptoms of myxœdema can be abolished by the prolonged administration of thyroid extract or thyroxine. (Fig. 623 and its legend should be carefully studied.)

2. **Cretinism.**—In this condition there is congenital absence or maldevelopment of the thyroid gland from causes which are not definitely known. Symptoms do not appear till the age of six months, possibly owing to the presence of sufficient thyroid substance in the mother's milk. The symptoms are as follows:

Fig. 624.—Cretin, aged 28, before treatment. Height, 34½ in. Murray's case, June 1895. (Swale Vincent, *Internal Secretion and Ductless Glands.* E. Arnold & Co.)

(i) The various milestones in the course of the child's development—holding the head up, sitting up, walking, speech—are reached much later than normal. Mental development may be very backward, and the child may be a complete idiot. There are anatomical changes in the central nervous system: nerve cells and fibres fail to appear or to mature; nerve tracts myelinate late or not at all.

(ii) Bony growth is retarded; the skeleton is small and stunted (Fig. 624).

(iii) The secondary sex characters do not develop.

(iv) The skin is thickened; the tongue is enlarged and protrudes from the mouth; deposits of fat appear, especially above the clavicles.

(v) The serum cholesterol rises (as in myxœdema, and is of similar diagnostic significance).

(vi) Creatine excretion in the urine is less in cretins than in normal children. The normal output on a meat-free diet (containing 2 g. of protein per kg. body weight) is 0·6–7·8 mg. daily; in cretins it falls to 0–3·8mg.

(vii) Constipation is commonly present.

Owing to lack of thyroxine, proper development of the body does not occur in the cretinous child, just as all bodily activities are depressed in the adult.

Considerable improvement can be obtained in cretins by means of thyroid therapy; if treatment is begun very early in life the child may grow up into a perfectly normal adult, but if treatment is delayed, less satisfactory development is obtained, particularly on the mental side.

3. **Simple Goitre** (Iodine-Deficiency Thyroid Hyperplasia).[1]—In

[1] British Medical Research Committee Memorandum No. 18, *Thyroid Enlargement and other Changes related to Mineral Content of Drinking Water,* 1948.

this condition there is enlargement of the thyroid, associated paradoxically with signs of *hypo*thyroidism. A survey of the distribution of the disease in different parts of the world shows that it is related to a relative or absolute *deficiency of iodine* in the water supply in relation to bodily needs. The normal iodine intake is 20–200 μg. daily. [Other environmental factors may also be involved, *e.g.* the presence in the diet of toxic agents called *goitrogens* *i.e.* substances that interfere with thyroxine synthesis in the thyroid, as explained on p. 988]. Simple goitre can be produced experimentally in animals on an iodine deficient diet (Figs. 625, 626).

The thyroid changes in simple goitre are produced as follows: because of the lack of dietary iodine or the toxic action of the goitrogens, thyroxine

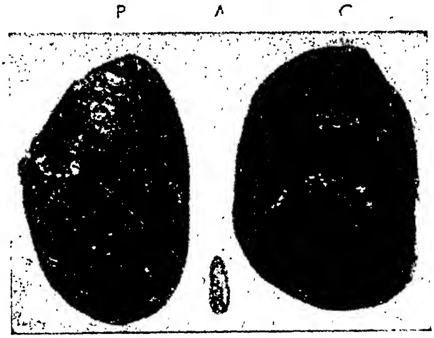


FIG. 625.—Experimental Production of Simple Goitres (Hyperplastic and Colloid) in Dogs. (Mellanby, *Nutrition and Disease*, 1934.)

A—Normal thyroid gland of dog. Weight 1 g.
B—Hyperplastic gland. Weight 100 g. resulting from iodine-deficient diet for 1½ years.
C—Colloid goitre. Weight 178 g. Iodine-deficient diet for 10 months, followed by 15 μg. of iodine daily for 8 months (1 μg. = 0·001 mg.).

formation is depressed. The anterior pituitary is stimulated to secrete excess thyrotrophin (p. 980); thyroid hyperplasia results, with the typical increase in thyroid weight, decrease in the amount of colloid in the vesicles and an increase in the height of the lining epithelial cells (Fig. 626, B). In this stage of hyperplasia the iodine content of the colloid of the human thyroid in simple goitre may be as low as 0·3 mg. per g. of dry thyroid (normal value is over 2 mg.); of the iodine present only a trace is in the form of thyroxine.

Sometimes the thyroid in a clinical case of simple goitre consists of vesicles which are not collapsed, but, on the contrary, enormously distended with colloid—*colloid goitre*; secondary degeneration may occur in such a gland, giving rise to cysts. Colloid goitres can also be produced experimentally; the procedure is first to produce a hyperplastic gland by means of severe iodine deprivation and then to add minute and quite inadequate amounts of iodine to the diet. In the experiment shown in Fig. 625, C a hyperplastic goitre was converted into a colloid goitre (with further increase in size from

100 to 178 g.) by adding 15 μg. of iodine daily to the diet (1 μg.=0·001 mg.). The amount of iodine administered was thus sufficient to produce colloid storage without restoring the gland to normal. Had 100 μg. of iodine been administered daily the gland *would* have become smaller and finally returned to normal.

Simple goitre can be prevented in most cases by remedying the iodine deficiency. In Switzerland since 1922 iodized salt has been available containing 1 in 200,000 iodine; no mother taking such salt during the last five months of pregnancy has given birth to a child with a goitre. In this connection it should be remembered that milk, meat, bread, cabbage, and vegetable oils contain no iodine; the daily consumption of sea fish (2 oz.)

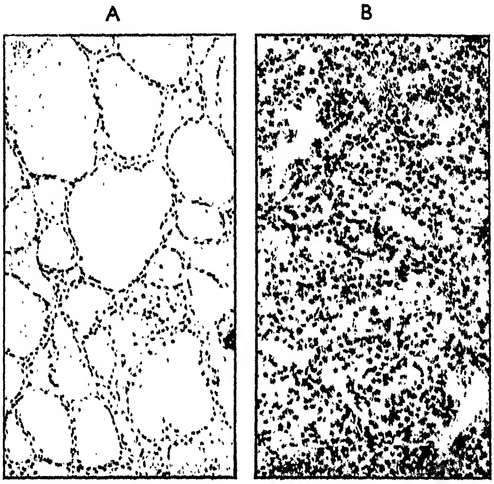


FIG. 626.—Relation of Iodine Content of Diet to Thyroid Structure (in Dog). (Mellanby, *Nutrition and Disease*, 1934.)

A—Thyroid on a diet containing cod-liver oil (containing iodine)—normal structure.
B—Thyroid on the same diet, except that the oil was pea-nut oil (contains no iodine)—marked hyperplasia.

and cod-liver oil (3 drms.), which are rich in iodine, forms an adequate prophylactic during pregnancy.[1]

The minimum iodine intake needed to prevent goitre probably varies with circumstances. In some conditions the bodily needs for food iodine may increase; unless additional amounts of iodine are then supplied, goitre may develop. Among such predisposing factors to simple goitre production are puberty, pregnancy, and certain infections. One mg. of I daily is sufficient to provide for all contingencies.

CHEMICALLY INDUCED THYROID HYPERPLASIA AND HYPOTHYROIDISM.—As already indicated the combination of hypothyroidism and thyroid hyperplasia, which is characteristic of simple goitre, can be induced experi-

[1] After 1951 all salt in Britain is to be "fortified" with KI or NaI in an amount equivalent to 15–30 parts of iodine per million parts of salt.

mentally by many agents administered in the diet (even when the iodine intake is adequate). Among the more important are certain *organic sulphur-*containing compounds (especially *thiouracil*), *aniline* derivatives (including the *sulphonamides* and amino-benzoic acid) and some plants of the Brassica family (*e.g.* rape seed). These *goitrogens* induce hypothyroidism by interfering with thyroxine synthesis in the gland. They have no peripheral neutralizing action on thyroxine so that their systemic effects on the animal as a whole are abolished by the coincident administration of thyroid hormone. The thyroid hyperplasia is due to excess thyrotrophin secretion (p. 980).

A

B

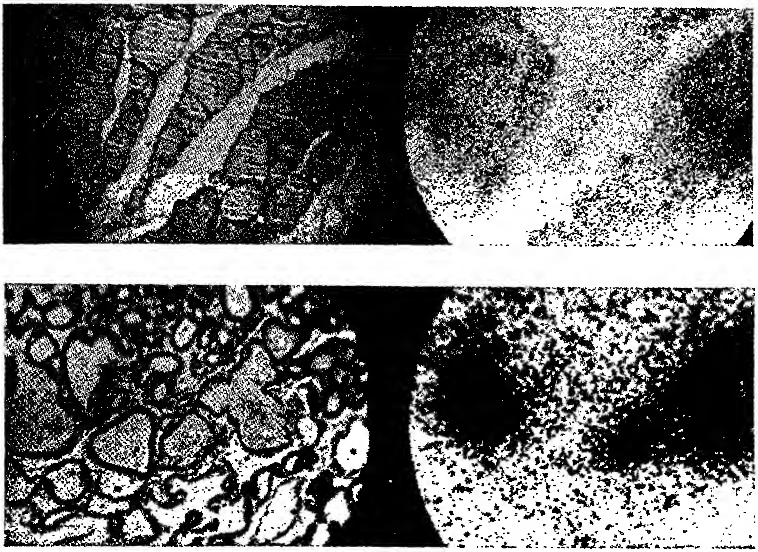


FIG. 627.—Structure and Radio-autograph of Thyroid in Normal and in Graves' Disease. (Hamilton, *Radiology*, 1942, *39*, 557.)

A. Normal thyroid. Left=Photomicrograph; Right=corresponding radio-autograph. Note that the presence of radio-active iodine in gland produces areas of darkening, thus demonstrating the distribution of thyroid hormone.

B. Hyperplastic thyroid. Left=Some of the large vesicles are filled with colloid, others are empty. Right=Radio-autograph shows that most of the radio-active iodine is in the former vesicles.

Hyperthyroidism (Graves' Disease, Exophthalmic Goitre).—In this disease there is excessive activity of the thyroid gland. The gland is usually only moderately enlarged; the vesicles vary in appearance, some are collapsed (Fig. 616, C), others are enlarged, and their epithelium is thrown into folds. Some of the large alveoli are empty of colloid, while others are filled with colloid rich in iodine (Fig. 627). The average iodine content of the colloid is very low, about 0·26 mg. per g. of dry thyroid (normal 2 mg.); the concentrations in the thyroid of all the three iodine fractions is reduced, the thyroxine being especially affected as it is being rapidly secreted into the blood. The plasma hormone iodine level is markedly increased and is probably the most reliable index of the degree of overactivity of the gland (Fig. 628).

It is interesting to note that in Graves' disease there is often great enlargement of the thymus and hyperplasia of the lymphoid tissues. Extravasations of lymphoycytes (lymphorrhages) are found in the skeletal muscles and especially in the eye muscles. Removal of the diseased thyroid leads to regression of the thymus. The significance of these findings is not known.

The initiating factor which causes the thyroid to become overactive has not been identified, but emotional stress acting via the hypothalamus on the anterior pituitary (p. 980) may sometimes play a part. Some cases of simple adenoma of the thyroid may, at some stage, begin to secrete excessively and produce symptoms of hyperthyroidism.

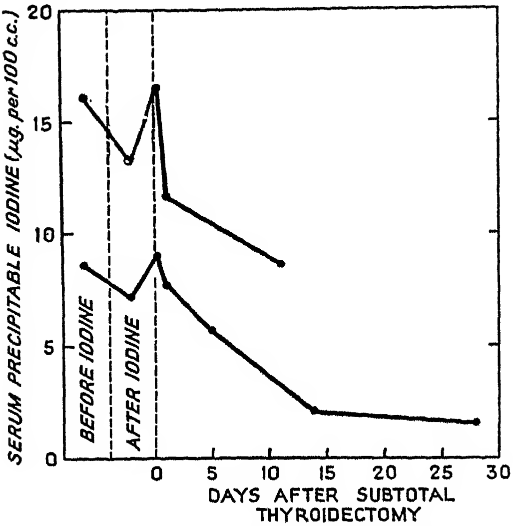


FIG. 628.—Changes in Serum Precipitable Iodine (Hormone Iodine) in cases of Hyperthyroidism before and after subtotal Thyroidectomy. (Winkler *et al.*, *J. clin. Investig.*, 1946, *25*, 407.)

The normal range of serum hormone iodine is 4–8 µg-%.
Administration of therapeutic doses of iodine temporarily lowered the level of serum hormone iodine. Administration of iodine was stopped on day of operation.
Removal of the thyroid further lowered the serum hormone iodine level. In the lower curve the level became subnormal.

The main changes in Graves' disease are as follows :

(1) METABOLIC RATE.—The cardinal symptom is the *increased basal metabolic rate,* which is the result of excessive secretion of thyroxine. The pulmonary ventilation, the oxygen consumption, the CO_2 output, the nitrogenous excretion in the urine are all increased as a consequence. The energy output at rest may in moderate cases be 50% above normal (*i.e.* B.M.R. +50) and in severe cases 100% above normal (B.M.R. +100) (Fig. 224, p. 377).

(2) BLOOD HORMONE LEVEL.—The level of hormone (precipitable) iodine in the serum is elevated to a degree which is proportional to the severity of the clinical disturbance (Fig. 628).

(3) CIRCULATION.—The rate of the heart is greatly increased (*e.g.* to 140 per minute at rest) ; the diastolic recovery period is thus greatly shortened. The cardiac output (and therefore the work of the heart) are likewise increased to supply larger amounts of oxygen to the overactive tissues. To increase heat loss and thus prevent a rise of body temperature resulting from the increased heat production, the cutaneous arterioles and capillaries are dilated and the skin is flushed and moist ; skin temperature is notably raised in the toes and to a less extent in the fingers. The blood pressure does not show constant changes.

Thyroxine accelerates the heart and increases its excitability partly by a direct action (p. 274) ; it is not known whether thyroxine in excess directly

damages the heart muscle. Signs of heart failure often supervene, and auricular fibrillation may develop.

(4) SPECIAL METABOLISM.—(i) There are signs of mild diabetes; the blood sugar may be raised (mainly because of increased conversion of glycogen to glucose in the liver) and *glycosuria* is frequently present. The glucose tolerance curve is raised, partly owing to more rapid glucose absorption from the bowel.

(ii) The serum cholesterol is lowered, *e.g.* from the average normal of 180 mg. down to 120 mg–%; the total blood fat concentration is likewise decreased.

(iii) In normal adults creatine is absent from the urine or present only in traces (normal range 0–60 mg./24 hr.); in Graves' disease considerable creatinuria develops (Fig. 632) associated with an elevation of the serum creatine level. Thyroxine apparently interferes with the retention of creatine in the muscles. Normally, if a test dose of 2·6 g. of creatine is ingested, about 70–80% is retained in the body (presumably in the muscles) the rest being excreted in the urine. In Graves' disease a much smaller fraction is retained and correspondingly more is excreted.

(iv) There is some rarefaction of the *bones* from mobilization of calcium salts and their excretion in the urine (serum calcium tends to fall).

(v) There is frequently depressed *liver* function as shown by the results of the galactose test (p. 830).

(5) A fine *tremor* of the voluntary muscles is present; its mode of production is unknown.

Rôle of Vitamin Lack.—Some of the changes in Graves' disease may be secondary to vitamin deficiency. Thus an increase in tissue metabolism leads to increased utilization and therefore to a greater need for the vitamins of the *B*-group. As the patient's appetite and food intake are poor, vitamin lack develops. In hyperthyroid animals it is possible to prevent the usual depletion of hepatic glycogen and the loss of weight, and also to reduce the pulse rate by giving the vitamin-*B* complex. It is claimed that clinically, pyridoxine (p. 1028) improves the muscular weakness, and that if vitamin-*D* is added to the diet the characteristic calcium loss from the bones may be prevented.

(6) EYE SIGNS IN GRAVES' DISEASE[1] (Fig. 629).—The principal eye changes are: (i) exophthalmos; (ii) retraction of the upper lids; (iii) weakness of the external eye muscles (ophthalmoplegia). All these eye signs are largely due to *excessive deposition of fat* at the back of the orbit and in the external eye muscles. The resulting increase in the volume of the retro- and peri-bulbar tissues pushes the eyeball forward, producing *exophthalmos*. The levator palpebrae superioris is very commonly infiltrated with fat to a considerable extent; this may be the cause of the spasm of this muscle which leads to the *retraction of the upper lid*. The fat content of the external eye muscles is always increased; in chronic cases these muscles show marked degenerative changes with disintegration, infiltration with lymphocytes and disappearance of many fibres (Fig. 630) which account well for the *muscular weakness*. It should be emphasized that these changes are not limited to

[1] Brain, *Lancet*, 1939, ii, 1217. Pochin, *Clin. Sci.*, 1939, *4*, 91. Rundle and Wilson, *Clin Sci.*, 1944–45, *5*, 17, 31, 51. Woods, *Medicine*, 1946, *25*, 113. Dobyns, *J. clin. Endocrin.*, 1950, *10*, 1202.

the eye muscles ; similar, though less marked, changes occur in all the skeletal muscles, and may account for the generalized muscular weakness. It is noteworthy that the accumulation of fat in the orbit and muscles occurs even in patients who are greatly wasted ; such hyperthyroid cases may have more

FIG. 629.—Exophthalmic Goitre. (Simpson, *Major Endocrine Disorders*, 2nd. edn., Oxford University Press, 1948)

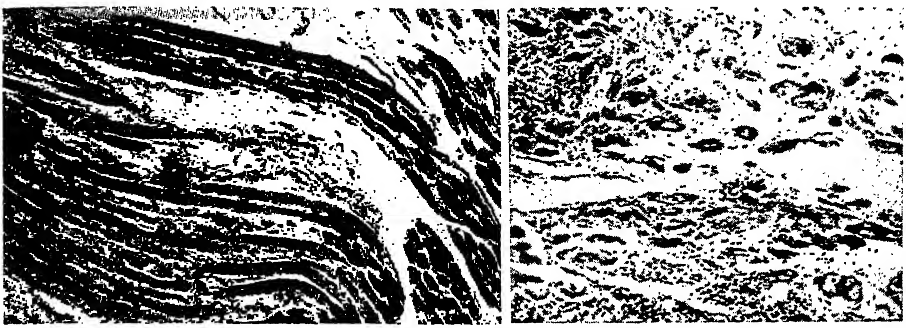
Note the eye changes.

fat in the sites named than obese (non-hyperthyroid) patients. In Graves disease there is thus a mysterious derangement of fat deposition in certai localities contributing to, or associated with, great muscular weakness.

Rôle of the Sympathetic Nerves.—In acute animal experiments exoph thalmos results from sympathetic stimulation which causes contraction c the retro-ocular muscle of Müller (p. 708). In man, however, this structur is vestigial and is unable to produce a bulging forward of the eye. As th

exophthalmos of Graves' disease persists after death it cannot be due to muscular contraction or to vascular engorgement. Section of the cervical sympathetic does not affect the exophthalmos or abolish the lid retraction; the latter is thus not due to overaction of the smooth muscle fibres in the upper lid; the striated fibres in the lid supplied by the third nerve must be the ones involved. When retraction of the upper lid is due to sympathetic overaction causing contraction of the smooth fibres of the superior tarsal muscle (p. 708), it is accompanied by a depression of the *lower lid*; in Graves' disease on the contrary, the lower lid is raised relative to the cornea. The pupils in Graves' disease are not dilated as would be the case if there were sympathetic over-action.

Rôle of Thyrotrophic Hormone.—As injection of thyrotrophin produces



A B

FIG. 630.—Changes in Eye Muscles in Exophthalmic Goitre. (Aird, *Ann. int. Med.*)

A. Normal extra-ocular muscle (man). Note well-formed, deeply stained long fibres.
B. Extra-ocular muscle in patient with severe exophthalmos (same magnification as in A). A few darkly staining fragments of muscle fibres are still visible, but the main features are œdema, extensive infiltration with round cells, degenerated muscle fragments and fibrosis.

exophthalmos in many species there has been a natural tendency to attribute the exophthalmos of Graves' disease to pituitary overaction (and to suggest, in addition, that the thyroid hyperplasia is due to the same cause). There are serious objections to this view:

(i) The orbital changes in Graves' disease (fat accumulation) are distinct from those produced by thyrotrophin in animals (œdema).

(ii) There is no proof of increased blood thyrotrophin in Graves' disease, but this may be due to the fact that the hyperplastic gland inactivates thyrotrophin rapidly (p. 980).

(iii) In conditions in which the blood thyrotrophin concentration is increased, *e.g.* in myxœdema, cretinism, simple iodine-deficiency goitre, and after administration of thiouracil and related "goitrogens," there are no eye changes.

(iv) In Graves' disease the eye changes follow an *independent course* from that of the hyperthyroid manifestations and are not related to them in time or severity. They are not necessarily relieved by thyroidectomy or other procedures which decrease the hyperthyroidism. Exophthalmos may be unilateral.

32

The remote cause of the eye changes in Graves' disease is therefore still obscure.

Treatment of Graves' Disease.—(1) OPERATIVE REMOVAL of sufficient thyroid tissue usually cures Graves' disease. The serum hormone level and the clinical state are the best guides to the amount of thyroid substance which must be taken away (Fig. 628).

(2) ACTION OF IODINE.—The administration of 10–70 mg. of iodine daily may lower the basal metabolic rate (B.M.R.), *e.g.* from +80 to +20 (Fig. 631).[1] The clinical condition improves for about 2 weeks, and the patient is better able to stand operative procedures; iodine medication also prevents the post-operative crisis which sometimes occurs with fatal results. Fig. 631 shows that when iodine treatment is continued for a longer time, the clinical condition begins to *deteriorate* and settles down to a state roughly midway between that originally present and the stage of maximum benefit. It is important to remember that the discontinuance of iodine treatment leads to an exacerbation of the symptoms.

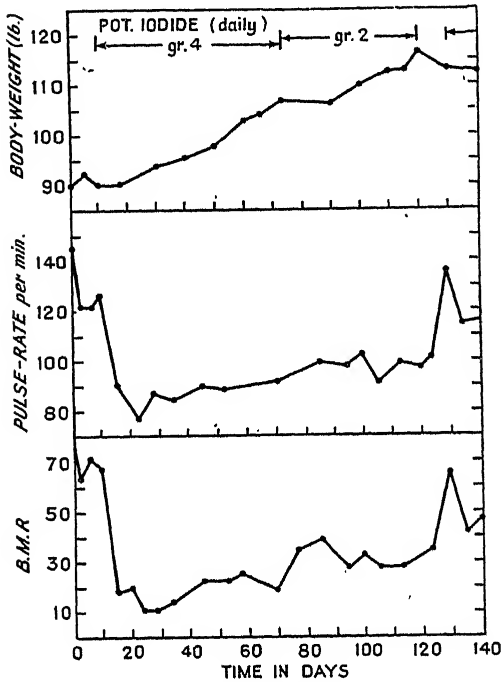


FIG. 631.—Effects of Therapeutic Administration of Iodine (as KI) in Graves' Disease. (After Mellanby, *Nutrition and Disease*, 1934.)

Record from above downwards:
Body-weight in lbs.
Pulse rate per minute.
B.M.R. as per cent. *above normal*.

Pot. Iodide was given first in doses of 4 grains daily. Note the increase in weight, the fall in pulse rate and the decrease in B.M.R. After 20 days, the pulse rate and B.M.R. began to rise again in spite of treatment; they stayed approximately at the new level when the dose was reduced to 2 grains daily.

When the iodide treatment was stopped, there was a sudden marked increase in pulse rate and B.M.R. They fell again on resuming treatment.

These large doses of iodine inhibit the secretion of thyroxine into the blood; the level of serum *hormone* iodine is decreased (Fig. 628). Iodine treatment depresses the thyroid epithelium (*i.e.* diminishes the hyperplasia), and promotes *storage of colloid* in the vesicles which become distended. The iodine content of the gland is increased, *e.g.* from 0·25 to 2·3 mg. of I per g. dry weight; all three forms of iodine are stored in larger amounts, including thyroxine. Iodine treatment does not bring about a complete cure as the gland is only partially, and temporarily, restored to a more normal, "resting" state.

[1] Therapeutically administered iodine acts by raising the blood inorganic iodide level (p. 982).

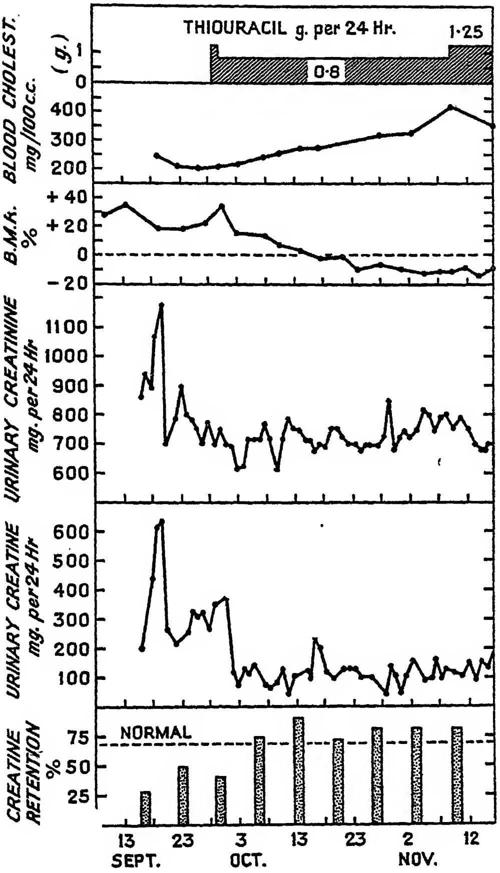


FIG. 632.—Effect of Thiouracil in case of Hyperthyroidism. (Barr and Shorr, *Ann. int. Med.*, 1945, *23*, 758.)

Records from above downwards :
Dose of thiouracil administered in g. per 24 hr.
Serum cholesterol in mg./100 cc.
B.M.R. % above or below normal (0=normal).
Urinary creatinine in mg./24 hr.
Urinary creatine in mg./24 hr.
Creatine retained as % of dose administered.
Administration of thiouracil restored metabolism to normal.

Note : Serum *cholesterol* rose (from 200 to 450 mg-%).
B.M.R. fell (from +40 to −10).
Urinary *creatinine* output fell (normal range 1200–1700 mg./24 hr.).
Urinary *creatine* : normal output 0–60 mg./24 hr. Note the high creatine output in hyperthyroidism (200–600 mg./24 hr.). After treatment the creatine output fell to 50–200 mg./24 hr.
Percentage creatine *retained* after a test dose (given by mouth) rose from 25–50% to over 75%.

(iii) ACTION OF THIOURACIL AND RELATED COMPOUNDS.—As already explained, many thio compounds inhibit thyroxine synthesis in the thyroid gland (p. 975). Some of these substances are very useful in the treatment of Graves' disease. Methyl thiouracil is most generally employed clinically. After a latent period of 1–2 weeks, the pulse rate and B.M.R. begin to fall, the body weight increases and the clinical state improves (Fig. 632). After the

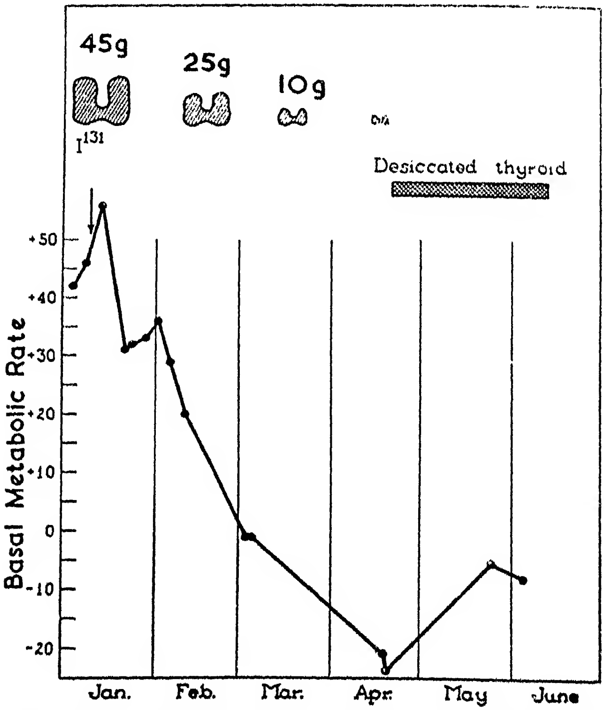


FIG. 633.—Therapeutic Action of Radio-Active Iodine ($^{131}I^*$) in Graves' Disease. (Kelsey *et al.*, *J. clin. Investig.*, 1949, *9*, 199.)

One dose of radio-active iodine was administered at the time indicated by the arrow. The thyroid decreased progressively in weight from 45 g. to 10 g. and finally almost completely disappeared. The B.M.R. fell from +50 to less than −20. The induced myxœdema was treated with desiccated thyroid.
$^{131}I^*$ gives off β and γ radiations.

maximal improvement has been attained, the patient is given small maintenance doses, which may completely control the symptoms and signs of hyperthyroidism; the thyroid enlargement, however, persists (or the gland may enlarge even further) and the eye changes are unaffected. The drug must be employed with care as agranulocytosis and other unpleasant side actions may develop. The thio compounds control the hyperthyroidism by *inhibiting thyroxine formation*. The fall of blood thyroxine may stimulate thyrotrophin formation and thus induce further thyroid hyperplasia. The *histological* appearance of the gland, therefore, remains that of *untreated Graves' disease*,

i.e. the vesicles are collapsed and colloid free, the lining cells are tall, *e.g.* 20 μ in height, and contain vesicular nuclei 10 μ in diameter; the gland is congested and friable and may be difficult to remove if operation becomes necessary. The thyroid cells may ultimately *degenerate*; if this occurs, further thiouracil treatment may become unnecessary. The general result, however, is to leave an enlarged vascular hyperplastic gland which has ceased making excess thyroxine because the drug has inhibited thyroxine formation.[1]

(iv) Administration of Radio-Iodine ($^{131}I^*$).—As already explained (p. 973) the thyroid "collects" I from the extracellular fluids, with the result that the thyroid inorganic iodide concentration may be hundreds of times greater than that in the blood or the tissue fluids generally. If suitable doses of radio-active I are administered to patients with Graves' disease, it accumulates in the thyroid in sufficient concentration to damage or destroy most of the alveoli without injuring the rest of the tissues of the body. Marked clinical improvement has been obtained in this way[2] (Fig. 633).

CALCIUM AND PHOSPHORUS METABOLISM

Calcium and Phosphorus Metabolism.—The absorption and excretion of calcium and phosphate, their level in the blood and in the other body fluids, and their behaviour in bone, are so closely inter-related that it is best to consider the metabolism of both substances together.

Calcium Metabolism.—Calcium in Food.—The chief sources of calcium in food are:

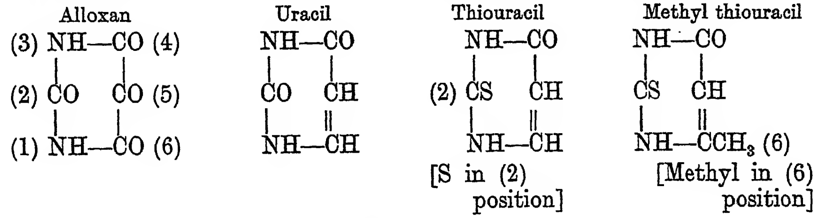
(i) Milk (120 mg./100 c.c. or about 660 mg./pint), and cheese (about 800 mg./100 g.). (*Human* milk contains only 30 mg./100 c.c.)

(ii) The concentration of calcium in the body fluids and in animal tissues (other than bone) is low; meat contains very little calcium (about 10 mg./100 g.). Most green vegetables are likewise poor in calcium and lose much of what they contain when boiled.

(iii) Flour contains 20–25 mg./100 g.; in Great Britain all flour has added to it 60 mg. of $CaCO_3$/100 g. with the result that bread has become an important source of calcium that can compensate for a deficiency of milk and cheese.[3]

Absorption of Calcium.—This takes place in the small intestine and is always incomplete; what is not absorbed is passed out in the fæces. The

[1] It is interesting to compare the formulæ of alloxan (which poisons the islets of Langerhans), and thiouracil and methyl thiouracil (which poison the thyroid).



[2] Kelsey *et al.*, *J. clin. Investig.*, 1949, *9*, 171.

[3] The special arrangements in force in Great Britain to ensure the larger calcium requirements of children and pregnant and lactating women are described on p. 1064.

(iv) Plasma inorganic phosphate level[1]: serum Ca generally varies inversely as plasma phosphate concentration, the product of the two (each expressed as mg–%) being normally 60 in children and less (about 40) in adults. The relationship is well illustrated in the Table below, which shows the effects of parathyroidectomy and subsequent injection of three doses of parathormone.

	Serum Ca in mg–%.	Plasma Inorg. Phosphate in mg–% of P.
Normal	11·0	5·5
After parathyroidectomy	5·2	10·9
Inject parathormone	6·7	8·8
” ”	11·2	5·7
” ”	15·1	2·0

There are however many exceptions to this generalization; thus in rickets (p. 1010) both the serum Ca and inorganic phosphate may be reduced.

Serum calcium is *lowered*:

(i) When there is reduced absorption of calcium from the intestine, *e.g.* in rickets and steatorrhœa, or when the calcium intake is very low.

(ii) In hypoparathyroidism.

(iii) In some forms of renal failure in which there is phosphate retention leading to a secondary fall of serum Ca.

Serum calcium is *raised*:

(i) In hyperparathyroidism.

(ii) In diseases in which abnormal plasma proteins occur (as stated above).

(iii) After administration of very large doses of vitamin-*D* (100,000–500,000 units per day).

(iv) After sudden immobilization, *e.g.* confinement to bed with extensive plaster casts as after fracture of the spine.

Excretion of Calcium.—(1) By the Bowel.—There is no doubt that calcium is *actively excreted* by the intestinal mucosa (possibly of the large intestine) from the blood into the lumen of the bowel. The evidence is as follows:

(i) If a calcium salt containing a trace of the radio-active isotope ($^{44}Ca^*$, half life 180 days) is injected intravenously in animals, 10% of the dose may be eliminated in the fæces in 10 days.

(ii) On a daily calcium intake in man of only 110 mg., as much as 200 mg. may be lost daily in the fæces.

The calcium passed out in the fæces is thus derived from two sources: (i) food calcium which has not been absorbed; (ii) body calcium which has been excreted from the blood by this route. No explanation has been offered to account for this excretion, which may continue even when there is calcium lack.

(2) By the Kidney.—The amount excreted in the urine varies widely, *e.g.* from 50 to 250 mg. per day; it is related to the amount absorbed and to the serum Ca level; if the latter is below 6–8 mg–% the urinary excretion is minimal; if it is raised above normal the urinary output correspondingly

[1] Inorganic phosphate concentrations are always expressed as mg. of P (per 100 cc.).

rises. The calcium output in the urine is raised in (i) hyperparathyroidism, (ii) acidosis, and (iii) hyperthyroidism (p. 991); in these conditions calcium is withdrawn ("mobilized") from the bones. With a normal calcium intake more calcium is lost in the fæces than in the urine.

Calcium Requirements.—To be in a state of calcium balance the calcium intake must equal that passed out in the fæces and urine. Calcium must be *retained* in growing children and during pregnancy; during lactation additional calcium is needed to make good the loss in the milk. The recommended *daily* calcium intake in various states (given below) is 50% above the probable minimal requirements, thus providing a margin of safety.

Adult: 12 mg./kg., or 750 mg. in all.

First 6 months: breast fed, 45 mg./kg.; artificially fed (cow's milk), 150 mg./kg. (the difference is due to the different degrees of calcium absorption).

6 *months–9 years*: 800–900 mg.

9–16 *years*: 1000–2000 mg. (1–2 g.).

Over 16 *years*: gradually decreasing to adult level.

Pregnancy: the calcium content of the fœtus at 28 weeks is 5 g.; at 40 weeks (p. 1059) it is 30 g. The calcium intake should be increased to allow for a calcium retention of 50 g. in all, to provide a generous margin of safety. The intake should be 1·5 g. in the first months of pregnancy rising to 3 g. daily in the last 3 or 4 months.

Lactation: over 3 g. are needed daily to allow for the calcium loss in the milk and the unexplained increased loss from the bowel. It should be emphasized that at the height of lactation a woman loses about ten times as much calcium daily in the milk as she lost to the fœtus daily during the last stages of pregnancy.

Functions of Calcium in the Body.—(1) BONE.—The inorganic material of bone consists of calcium salts of phosphate (85%) and carbonate which are deposited in a protein matrix; there is a nucleus of $Ca_3(PO_4)_2$ on to which is absorbed chiefly $CaCO_3$ but also $CaHPO_4$ and $Ca(OH)_2$. Half the weight of a bone is calcium salt; 97% of the body calcium is found in the bones. The calcium of bone is *constantly undergoing exchanges with the body fluids.* Thus, if a calcium salt containing a trace of the radio-active isotope $^{44}Ca^*$ is injected intravenously 86% of the dose is taken up by the bones in 100 minutes; about 5% of this skeletal $^{44}Ca^*$ is given up in 10 days and replaced by a corresponding amount of calcium which has been taken up from the body fluids. The entire calcium of the bones might thus be "turned-over" in some 200 days. The special affinity of bone for calcium is shown by the fact that it takes up calcium about 130 times as rapidly as skeletal muscle. As the salts of bone are derived from the body fluids, satisfactory calcification depends on a normal serum concentration of Ca and phosphate ions. Other important factors are:

(i) *Alkaline phosphatase*, which is present in the centre of ossification in the hypertrophied cartilage cells and in the osteoblasts; it acts on the hexose phosphate brought in the blood and releases inorganic phosphate, thus raising its local concentration. The solubility product for calcium phosphate is consequently exceeded, with the result that insoluble calcium phosphate is precipitated out. The mode of precipitation of calcium carbonate is unknown.

(ii) *Vitamin-D*, which in addition to maintaining the normal serum Ca and inorganic phosphate levels by promoting absorption in the intestine also exerts an obscure *direct* action on the bone itself.

Other functions of calcium are discussed elsewhere as indicated below.

(2) Excitability of nerve fibres and nerve centres (p. 1004).

(3) Relation to clotting of blood (p. 140) and milk (p. 776).

(4) Relation to contraction of heart muscle (p. 236).

Phosphorus Metabolism.[1]—PHOSPHORUS IN FOOD.—The chief food sources are set out below.

(i) Milk (90 mg./100 g. or about 500 mg./pint) and cheese (about 700 mg./100 g.) mostly as inorganic phosphate and partly in the phosphoprotein caseinogen.

(ii) P is found in animal tissues in: (*a*) nucleoproteins and nucleotides (p. 878); (*b*) phosphatides, *e.g.* lecithin (p. 861); (*c*) phosphoproteins (caseinogen, ovalbumin of egg); (*d*) combined with coenzymes, *e.g.* flavin, (p. 1028); coenzyme-I (p. 840); (*e*) as inorganic phosphate. Skeletal muscle contains creatine phosphate (p. 429), adenosine triphosphate (Fig. 553), and hexose phosphate and its triose derivatives.

ABSORPTION OF PHOSPHORUS.—Phosphorus is always absorbed from the small intestine as inorganic phosphate; thus the organically bound phosphate must first undergo appropriate digestion in the intestine to release the inorganic phosphate. Phosphate absorption is facilitated by the formation of the more *soluble salts* and by the presence of *vitamin-D*.

PHOSPHORUS IN BLOOD.—The red cells contain large amounts of *organically* bound P (85 mg./100 c.c.) in phosphate esters and phosphatides. The plasma P is mostly in the form of *inorganic phosphate*. The normal serum inorganic phosphate concentration is 5–6·5 mg–% in infants and 3–4·5 mg–% in adults; this phosphate is mainly in the form of B_2HPO_4.

Serum phosphate is *reduced* in rickets, in the related condition of osteomalacia, in defects of the intestinal mucosa (*e.g.* sprue) and in uncomplicated hyperparathyroidism; it also falls temporarily, after a carbohydrate meal or an injection of insulin.

Serum phosphate is *raised* in renal failure, acromegaly, after injection of pure growth hormone, and in simple hypoparathyroidism. The inverse relationship commonly (but not invariably) found between serum phosphate and Ca concentrations was considered on p. 999.

EXCRETION OF PHOSPHORUS.—The P passed out in the fæces is food P which has not been absorbed. On an average 60% of the total P loss is in the urine. Both parathormone and thyroxine increase the elimination of phosphate in the urine (p. 1003). The rôle of the phosphate in the urine is discussed on p. 94.

PHOSPHORUS REQUIREMENTS.—The adult requirement is 1·4 g. daily; on any reasonable diet the P supply is sufficient for all needs.

FUNCTIONS OF PHOSPHORUS IN THE BODY.—(1) *Bone* (see p. 1000).

(2) *Metabolism.*—Organic phosphate is present in large amounts in all tissue cells in many forms (cf. Fig. 2). It is concerned with numerous metabolic changes in the body in connection with carbohydrate (p. 844), skeletal muscle (p. 429), and fat metabolism (p. 864); it is concerned in the action

[1] Phosphorus concentrations are expressed in terms of mg. of *P* (and *not of* phosphate) per 100 c.c. or 100 g.

of many coenzymes (p. 840). The importance of high energy phosphate bonds in energy transfer is discussed on p. 842.

(3) Regulation of *H-ion concentration* of blood and urine (see p. 94).

(4) In some circumstances inorganic phosphate is released from the organic phosphate in the tissue cells, passes into the blood and serves as a source of urinary phosphate (p. 95).

Phosphatase Activity.—Phosphatase hydrolyzes monophosphoric esters releasing inorganic phosphate. Two types of phosphatase are found in the body, the *alkaline* and the *acid*.

(1) ALKALINE PHOSPHATASE,[1] which acts in an alkaline medium, is present in all tissues and body fluids: the highest concentration is found in actively ossifying cartilage (centres of ossification, growing epiphyses) and periosteum, intestinal mucosa, kidney, and liver. Alkaline phosphatase is concerned with many chemical processes, *e.g.* (i) precipitation of calcium phosphate in bone (p. 1000); (ii) hydrolysis of phosphate esters in the intestine leading to phosphate absorption; (iii) hydrolysis of phosphate esters in the kidney which have been formed presumably in connection with the tubular reabsorption of glucose and other constituents; (iv) hydrolysis of glucose-6-phosphate in the liver releasing glucose into the blood (p. 845); (v) possibly with the synthesis of tissue protein (p. 935).

The normal *serum* alkaline phosphatase (in Bodansky units per 100 c.c.) is 1–4 units in adults and 5–14 units in children; it is derived chiefly from the *bones*. The level is *increased* in the following conditions:

(i) Diseases of *bone*, presumably because of increased local formation of phosphatase, *e.g.* in active rickets (to 20–190 units), osteitis deformans (Paget's disease) (to 15–125 units), hyperparathyroidism (to 20–40 units), and secondary deposits in bone from a primary carcinoma of prostate (p. 1116). It is also increased by injection of pituitary growth hormone.

(ii) Disorders of the *liver* and biliary tract (obstructive and toxic jaundice): the explanation is obscure. The rise has been attributed to failure of the liver to secrete the enzyme into the bile or to regurgitation of the enzyme from the bile into the blood.

(2) ACID PHOSPHATASE, which acts in an acid medium (*p*H 5·5), is formed in small amounts in many organs: the normal serum concentration is less than 3 units per 100 c.c. The prostate after puberty has a high acid phosphatase content (p. 1115); in cases of carcinoma of the prostate with secondary deposits in bones and lymphoid tissue, the serum acid phosphatase is usually raised. A value exceeding 10 units per 100 c.c. is abnormal (p. 1116).

THE PARATHYROID GLANDS. VITAMIN-*D*

The Parathyroid Glands.[2]—The parathyroid glands are normally four in number. They consist of masses or columns of cells, with numerous wide vascular channels between them. The cells are of two types: (i) *principal* or non-granular cells with faintly staining nuclei; (ii) a few *eosinophil* cells with well-marked nuclei; these latter cells do not appear until puberty. The significance of this histological differentiation is unknown.

[1] Morris and Peden, *Quart. J. Med.*, 1937, *6*, 211.

[2] Albright and Reifenstein, *Parathyroid Glands and Metabolic Bone Diseases*, Baltimore, 1948.

Parathyroid Hormone, Parathormone.—The active principle of the parathyroids is called *parathormone*. It is prepared by acid extraction of fresh parathyroid glands; it is freely soluble in water, and is usually administered subcutaneously or intravenously; it is effective when given by the mouth in large doses. The results of deficiency or excess of parathormone are well established and are described in detail below. The most striking and obvious effect of parathormone is on the level of *serum calcium*, but the *primary* site and exact mode of action of the hormone are, however, still obscure. The following suggestions have been made:

(1) The primary action of the hormone is to regulate the *excretion of inorganic phosphate in the urine.*

(i) The first observable effect of injecting parathormone in a case of clinical hypoparathyroidism or in normal subjects is to *increase* phosphate excretion in the urine (Fig. 634).

The subsequent course of events is thought to be as follows:

(ii) As a result the serum phosphate level falls; fæcal excretion of P is unaltered.

(iii) Because of the reciprocal relationship between serum Ca and serum phosphate concentrations, the fall of serum P leads secondarily to an increase in serum Ca; the latter may rise to 15 mg-% or more. This hypercalcæmia develops somewhat gradually and is associated with a flow of calcium out of the bones into the blood.

(iv) The hypercalcæmia leads to an increase in calcium excretion by the kidney.

The net result is rarefaction

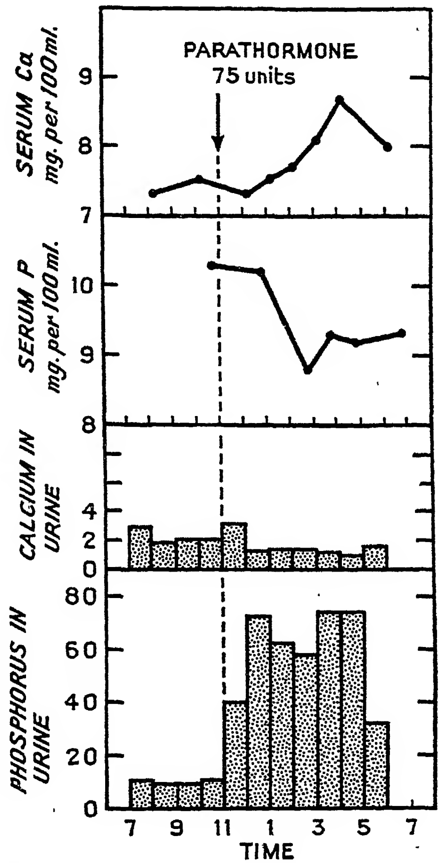


FIG. 634.—Effect of injection of Parathormone on Urinary Phosphate and Calcium Excretion and on Serum P and Ca. (After Albright and Ellsworth, *J. clin. Investig.*, 1929, 1, 183.)

Case of Idiopathic Hypoparathyroidism.
At the time marked by the arrow inject 75 units of parathormone.
Time in hours.
Ordinates above: Serum Ca and P in mg/100 c.c. Note high initial serum P (10·5) and low serum Ca (7·3).
Ordinates below: urinary excretion of Ca and P in mg/hr.
Note that the increased excretion of P in the urine produces the initial fall of serum P. The rise of serum Ca is secondary. The urinary excretion of Ca was increased later (not shown in the Fig.) as a result of the hypercalcæmia.

[illegible] and the elimination of the bone calcium in the [illegible] in the fæces is unaltered.

[illegible] excretion of phosphate in the urine could be most [illegible] increased reabsorption of phosphate by the renal [illegible] action of parathormone on the renal tubules has yet [illegible] has suggested that parathormone "affects [illegible] the body fluids in such a way as to make it more [illegible]"; this phrasing suggests that parathormone [illegible] electrolyte equilibria of the body," the other changes [illegible] still remain in the dark as to what is the precise

[illegible] have *a direct action on bone*. By stimulating [illegible] instance, more calcium would be liberated from [illegible] produce a rise in serum calcium concentration which [illegible] initial changes in renal activity. That this may [illegible] the fact that parathormone causes a rise in serum [illegible] bilateral nephrectomy. Alternatively, parathormone [illegible] *solubility of calcium in plasma* and so disturb the electrolyte [illegible] extracellular fluid and bones, leading to an outflow [illegible]

[illegible] may have *multiple* sites of action, *e.g.* on the kidney, [illegible]" and the bones.

Extirpation of the Parathyroids.—The main results in animals and [illegible]

[illegible] Parathormone lack leads to decreased excretion of phosphate [illegible] as little as 2% of the preoperative level.

As a result of the urinary changes the serum phosphate [illegible] mg-%; because of the inverse relationship between [illegible] phosphate, serum Ca falls, *e.g.* from 11 to 5·2 mg-%; [illegible] the *diffusible* (and ionized) Ca.

[illegible] All the changes in the nervous system (peripheral [illegible] after parathyroidectomy are due to the associated [illegible] The effects of a fall of Ca^{++} concentration must therefore be [illegible] detail.

[illegible] If a nerve fibre is bathed in fluid containing a [illegible] concentration, its excitability is increased: it responds to [illegible] threshold to the normal fibre; it may respond to a *single* [illegible] repetitive discharge; it may generate a *spontaneous* train of [illegible] the excitability of ganglion cells to acetylcholine is [illegible] Ca concentration (Fig. 319). It is suggested that a low [illegible] the "accommodation" of nerve, *i.e.* the process by [illegible] the action of the exciting stimulus (p. 494). The excita[illegible] nervous system is also enhanced by hypocalcæmia. The [illegible] are potentiated by ischæmia.

[illegible] The following observations can be made in man:[2] [illegible] wrapped round the upper arm of a *normal* person and inflated [illegible] artery; the nerve trunks which are *directly* com[illegible] are thus deprived of their blood supply. This ischæmia

[illegible] Pitts, *Amer. J. Physiol.*, 1948, *155*, 42.
[illegible] *Arch. Neurol. Psychiat.*, 1948, *60*, 140.

initially temporarily enhances the excitability of the affected nerve fibres, particularly those which have run the longest course from the periphery. After a minute's compression the fibres from the skin of the hand begin to discharge spontaneously; these afferent impulses on reaching the brain give rise to a tingling sensation which is referred to the hand. [If the ischæmia is maintained the excitability of the nerves under the cuff declines and the tingling disappears; later still the nerve fibres are paralysed and anæsthesia and weakness set in (p. 752).]

(ii) If *hypocalcæmia* is present owing to parathyroidectomy, the procedure described above produces more striking results. Spontaneous discharge of the affected nerve fibres sets in sooner and affects the motor as well as the sensory fibres. The muscles of the hand contract, their pattern of activity depending on the number of motor nerve fibres affected and their discharge rate. If a few motor fibres discharge at a low rate, irregular muscle twitchings occur; if many fibres discharge at a higher rate powerful muscular spasms develop; the position taken up by the hand depends on the relative strength of contraction of the opposing muscle groups.

In any clinical condition in which the diffusible (ionized) serum Ca is lowered, *hyperexcitability* of the peripheral nerves and of the central nervous system develops; the resulting state, known as *tetany*, is described below.

(4) Clinical Manifestations of Tetany.—(i) *Carpopedal Spasm.* (*Trousseau's Sign.*)—Pressure is applied to a limb to compress the local nerves. As explained above, the combination of ischæmia and hypocalcæmia causes both the sensory and motor fibres affected to discharge spontaneously. The hand generally adopts the "obstetric" position; the metacarpo-phalangeal joints are flexed, the fingers are extended, the thumb is drawn on to the palm, the wrist and elbow are flexed. In the case of the lower limb the toes are plantar-flexed, and the feet are drawn up. There is tingling in the distal part of the limb.

(ii) *Facial Irritability* (*Chvostek's Sign*).—Because of the heightened excitability of the nerves to mechanical stimulation, tapping over the facial nerve after its exit from the stylo-mastoid foramen results in contraction of the facial muscles.

(iii) *Laryngismus Stridulus.*—Without warning, the laryngeal muscles contract suddenly, and the glottis is closed; this may be the result of a spontaneous discharge of the motor nerve fibres. No air can enter the chest, and progressive cyanosis develops. After a variable period the spasm relaxes, and air enters with a "crowing" sound.

(iv) There is increased excitability of the motor nerves to the galvanic current. Normally a kathodal opening contraction (k.o.c.) can be obtained with a current of not less than 6 milliamperes. When contraction results from a current below 5 milliamperes, increased galvanic excitability is present. In tetany, 0·6 milliamperes may be sufficient to produce a contraction. A *slowly rising current* which is ineffective when applied to normal fibres because of the development of accommodation readily stimulates the nerve (*e.g.* ulnar) in states of hypocalcæmia, producing tingling and muscle spasm.

(v) Generalized convulsions may occur owing to enhanced central nervous system excitability.

Parathyroidectomy in man may be carried out accidentally in the course

of an extensive excision of the thyroid gland or as a deliberate therapeutic measure in cases of parathyroid tumour or of parathyroid hyperplasia.

TREATMENT OF PARATHYROID DEFICIENCY.—(i) The symptoms of tetany can be temporarily relieved by injecting parathormone which raises the serum calcium [1]; but as this calcium is withdrawn from the bones, such treatment is inadvisable for more than short periods in accidental parathyroidectomy, and is especially undesirable in patients operated on for parathyroid hyperplasia in whom the bones are already dangerously weakened.

FIG. 635.—X-ray Appearance of Bones in Clinical Hyperparathyroidism.

Case of generalized osteitis fibrosa (parathyroid adenoma). *Left:* X-ray of right forearm of patient (P). *Right:* that of normal limb (N) simultaneously taken as control. Note general and marked rarefaction and deformity of bones. (Hunter, *Brit. J. Surg.*, 1932, *19*.)

(ii) The serum Ca may be raised temporarily by intravenous injection of soluble calcium salts.

(iii) The long term treatment consists in promoting greater absorption of calcium from the intestine. Large doses of calcium (3 g. daily) are given by mouth (*e.g.* 7 g. of $CaCO_3$ or 22·5 g. of calcium lactate daily); in addition very large doses of vitamin-*D* are given (2·5 mg. daily = 100,000 international units (i.u.) = 100 times the therapeutic dose in rickets.) A synthetic steroid, *dihydrotachysterol* (AT10) given in doses of 0·1 c.c. daily acts like vitamin-*D* and promotes calcium absorption from the gut.

HYPERCALCÆMIA.—With overdoses of parathormone, hypercalcæmia and decalcification of bones occurs. Finally severe toxic symptoms develop.

[1] See Table on p. 999, for changes in serum Ca and phosphate.

The serum Ca rises to a critically high level; the usual inverse Ca/P relationship is not observed and the serum phosphate is also high. *Renal failure* occurs: the blood urea and non-protein-N are raised. There is loss of appetite, diarrhœa, dullness, drowsiness, and general muscular flaccidity; circulatory failure and coma ultimately develop.

Repeated injections of parathormone may, however, lead to the establishment of a non-reactive state in which the hormone no longer affects calcium or phosphorus metabolism. This is presumably because parathormone is a protein and has been prepared from the glands of a different animal from the one to which it is given experimentally. After repeated injections, antibodies to the foreign protein are probably formed which react with the injected material and interfere with its action. There is no evidence of the development of such a tolerance to "endogenous" parathormone, *i.e.* a patient with hyperparathyroidism always responds to the excessive amounts of hormone which he secretes.

Causes of Clinical Tetany.—Tetany occurs owing to (i) hypocalcæmia, (ii) alkalæmia.

(1) Hypocalcæmia.—Tetany from hypocalcæmia occurs (i) after parathyroidectomy (p. 1005), (ii) in rickets (p. 1010) and in osteomalacia (p. 1012), (iii) in renal failure with phosphate retention (p. 77). As already emphasized, it is the fall in *diffusible* serum Ca which affects nerve tissue; alterations in the amount of calcium combined with plasma protein are without action.

(2) Alkalæmia.—Alkalæmia produces typical tetany without altering the level of diffusible serum Ca; it can be shown that the change in blood reaction *directly* increases the excitability of nerve tissues. The symptoms are relieved by raising the H^+ ion concentration (*e.g.* by inhaling CO_2-rich mixtures or administering acidifying salts such as NH_4Cl).

Clinically, tetany from alkalæmia occurs:

(i) From excessive ingestion of alkali, *e.g.* sodium bicarbonate, as in the intensive treatment of peptic ulcer.

(ii) In association with profuse vomiting, as in pyloric obstruction (p. 105).

(iii) Following excessive pulmonary ventilation, *e.g.* voluntary hyperpnœa (p. 408), or the hyperpnœa of nervous disorders like encephalitis lethargica (see legend to Fig. 56, p. 99).

Control of Parathyroid Secretion.—The glands are *not* controlled by the nervous system or by the anterior pituitary. It is probable that the secretion of the parathyroid gland is regulated by the calcium concentration of the blood reaching it.

(i) Intravenous injection of 40 mg. of Na oxalate in a dog initially lowers the serum Ca (by precipitating calcium oxalate); rapid recovery of the serum Ca occurs in normal animals, but no recovery takes place in parathyroidectomized animals.

(ii) If decalcified blood is perfused through the parathyroids (and thyroid) of a dog and the venous outflow is infused into another animal it raises the serum calcium in the recipient by, *e.g.* 1·5 or even 4·5 mg-%.

(iii) It is claimed that in several conditions associated with lowered serum Ca there is diffuse hypertrophy of the parathyroid glands.

Clinical Hyperparathyroidism.[1]—This condition may result from *diffuse* hyperplasia of all the parathyroids or from a *localized* tumour

[1] Black, *Surg. Gynec. Obstet.*, 1948, *87*, 172.

(parathyroid adenoma, or, rarely, carcinoma) in one of them. The effects of the resulting over-secretion of parathormone on the urine, blood, and bones are those described on p. 1003 (effects of injection of parathormone). Briefly, there is a rise of serum Ca (even up to 29 mg-%) and a fall of serum phosphate; withdrawal of Ca from the bones; increased Ca excretion in the urine; initially increased and later decreased phosphate excretion. When the hypercalcæmia is severe it produces the toxic effects described on p. 1006.

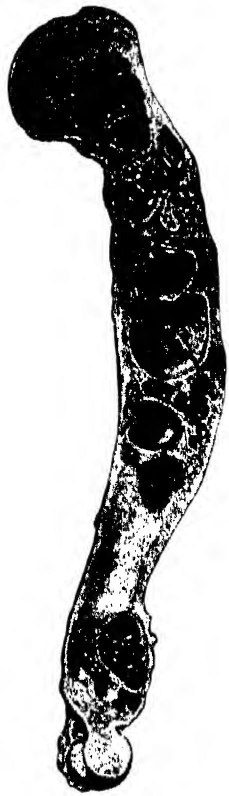


Fig. 636.—Bony Changes in Hyperparathyroidism.

Case of generalized osteitis fibrosa (parathyroid adenoma). Sagittal section of right humerus from same patient as Fig. 635. In the shaft are many dark areas representing cysts which were either empty or filled with deep red jelly. Note the great deformity of the shaft. (Hunter, *Brit. J. Surg.*, 1932, *19*.)

The variable changes which may occur in the bones are summarized below.

(i) There may be generalized rarefaction of the bones (osteoporosis) (Fig. 635).

(ii) The rarefaction may be more severe in certain bones, leading to formation of cysts; when many of these cysts are present, the condition is called *osteitis fibrosa cystica* (Fig. 636).

(iii) If calcium absorption from the intestine and deposition of fresh Ca in the bones keeps pace with the drain of calcium out of the bones the skeleton may remain intact.

The increased excretion of Ca in the urine (which is secondary to the hypercalcæmia) may lead to formation of renal calculi. In addition to the disabilities resulting from bony deformation (*e.g.* fractures) acute attacks of parathyroid poisoning may occur, characterized by restlessness, tachycardia, prostration and vomiting.

In cases of parathyroid adenoma, removal of the tumour leads to a rapid fall of serum Ca, commonly to subnormal levels, with resulting tetany; serum phosphate may rise (Fig. 637). The fall in serum Ca may set in within a few hours. The blood changes are well illustrated in the Table on p. 1009 (from C. E. Dent); in this case no rise of serum P occurred during the first 10 days after the operation.

The post-operative hypoparathyroidism is treated as described on p. 1006; in successful cases the bones are ultimately effectively recalcified.

Renal Rickets.—In some cases of renal disease in children, secondary parathyroid hyperplasia occurs, leading to progressive decalcification of the skeleton.

Vitamin-*D*.[1]—Vitamin-*D* increases calcium and phosphate absorption from the intestine and also promotes calcification by a direct action on bone;

[1] Hess, *Rickets, Osteomalacia and Tetany*, London, 1930. *Vitamins*, M.R.C., Spec. Rep., No. 167, 1932. Bricknell and Prescott, *Vitamins in Medicine*, London, 2nd end., 1947.

it prevents and cures rickets. The vitamin is found mainly in *fish-liver oils* (*e.g.* cod- or halibut-liver oil) and in yolk of egg; it is present to a small

Date.	Serum Ca in Mg–%.	Serum P in Mg–%.
16 June	15·0	1·5
27 „	14·2	1·5
3 July (day of operation)	14·9	—
4 „	13·8	1·1
6 „	11·6	—
9 „	9·6	—
10 „	9·4	1·3
12 „	10·3	—
14 „	9·8	1·2

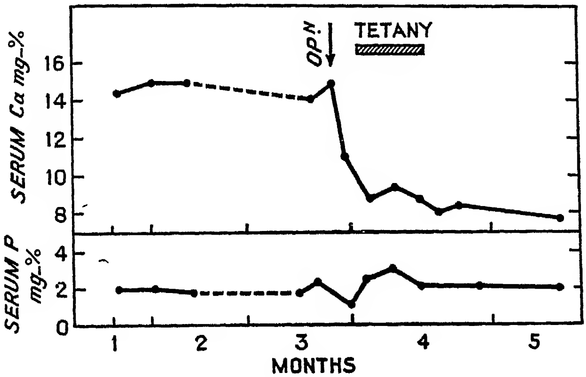


FIG. 637.—Blood Changes before and after Removal of Parathyroid Tumour.

The ordinates represent serum calcium and inorganic phosphate in mg–%.
The blood was examined for 3 months before and 2 months after operation. The parathyroid tumour was removed at the point marked by the arrow. Note the subsequent fall of serum calcium to below the normal value of 10 mg–%. Latent tetany developed. (Hunter, *Brit. J. Surg.*, 1932, *19.*)

degree in *animal fats* (*e.g.* beef suet, milk), and is practically *absent from vegetable oils* (*e.g.* olive oil) and vegetable margarines (British margarine, however, is now fortified with 60 i.u. per oz.). There are two main types of vitamin-*D*: *vitamin-D_2* (*calciferol*) and *vitamin-D_3*. Vitamin-D_2 is formed by ultraviolet irradiation of ergosterol, a substance found only in lower plants such as yeast, and fungi (*e.g.* ergot, from which it is named). *Vitamin-D_3* is the naturally occurring substance found in egg-yolk, butter, and fish liver oils. It is formed from 7-dehydrocholesterol present at the surface of the skin; activation by ultraviolet light opens ring B to produce vitamin-D_3 (Fig. 638). The international unit (i.u.) of vitamin-*D* is the activity of 0·025 μg of crystalline vitamin-D_2 (calciferol): *i.e.* 1 mg. of -D_2=40,000 i.u.

The vitamin-*D* content of certain foodstuffs is as follows: cod-liver oil average 500 i.u. per drachm; halibut-liver oil, 50 i.u. per drop; fresh milk,

usually less than 50 i.u per pint; butter 10–100 i.u. per ounce; eggs, 60 i.u. per yolk of 20 g. The vitamin-*D* content of milk, butter, and eggs depends on the vitamin content of the diet and on the season, being higher in the summer than in the winter. Milk, of course, is also rich in Ca and P. The vitamin-*D* content of the common foodstuffs is set out in the tables on pp. 1052 *et seq*.

Requirements.—No one knows how much vitamin-*D* is needed by adults; 250 i.u. daily are undoubtedly adequate. Pregnant and lactating women have larger requirements, as they must supply calcium for the fœtal skeleton during intra-uterine life, and for the milk (p. 1059). Growing children need extra supplies of calcium to form new bone. Both classes (*i.e.* mothers and children) need 500 i.u. of vitamin-*D* daily to promote the necessary calcium absorption. Some vitamin-*D* is synthesized in the skin (p. 1011). 1 to 2 pints of milk and an egg daily, plenty of butter, fortified margarine, and "fat" fish, may supply 300 i.u. If the diet is inadequate (as it is in wartime and in most countries in peacetime), and even if it is adequate, pregnant and nursing women, and babies after weaning and up to the age of seven years, should receive supplements of vitamin-*D* concentrates (*e.g.* cod- or halibut-liver oil) (cf. p. 1064).

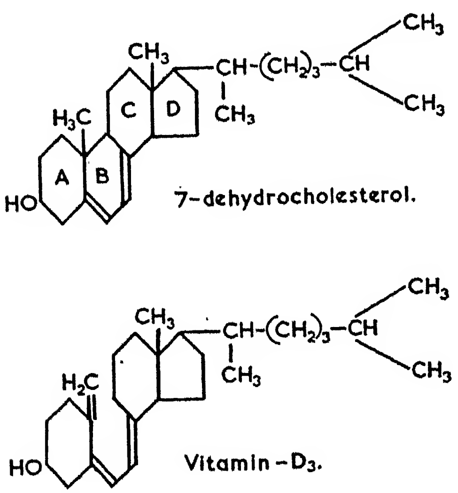


Fig. 638.—Formation of Vitamin-D_3 from 7-Dehydrocholesterol.

Rickets.—This disease of children is characterized by bones which are soft from deficient deposition of calcium salts and are therefore easily bent under the weight of the body, so that serious deformities may develop. Further, the process of ossification at the epiphyseal line takes place in an abnormal manner. Normally the epiphyseal line is a well-defined narrow strip of cartilage 2 mm. deep, behind which regular ossification is proceeding. In rickets the epiphyseal line forms a wide, irregular band, which can be felt as a marked projection on the surface. Normally the older cartilage cells degenerate and disappear, leaving many spaces into which the blood vessels and osteoblasts of the shaft can penetrate. In rickets this apparently essential preliminary degeneration does not occur and so ossification is retarded. The cartilage cells persist and go on multiplying, giving rise to a characteristic broad irregular cartilaginous zone. In addition, the matrix between the cartilage cells and that of the new bone itself does not become adequately impregnated with lime salts, accounting for the softness of the bones. *The level of serum calcium or of phosphate or of both is lowered*; in active rickets the product of serum Ca $\times$ serum P (each expressed in mg./100 c.c.) is 30 (normal=60). Serum alkaline phosphatase is commonly raised.

Rickets usually sets in about the sixth month of life, and its intensity is directly related to the rapidity of bodily growth. The disease may last for several years with final healing. There is conclusive evidence that the disease is due essentially to lack of vitamin-*D*.

(i) A condition identical with human rickets is induced in puppies if the diet is deficient in the foodstuffs containing vitamin-*D*. The disease is prevented or cured by the administration of cod-liver oil or egg-yolk. In puppies over 3 months old it is very difficult to induce rickets. Similarly, human rickets tends to heal after the age of 2 years, even if the diet and environment remain unaltered.

When a diet has a barely minimum vitamin-*D* content it is found that rickets may be precipitated by increasing the cereal content of the food, *e.g.* by giving large amounts of bread and especially oatmeal. The harmful agent in cereals is *phytic acid*, which interferes with the absorption of calcium by precipitating insoluble calcium phytate.

(ii) An examination of the diets of *children* proves that rickets is associated similarly with a deficiency of foods containing vitamin-*D*, with an excess of cereals, and a deficient calcium intake. Prophylaxis and cure result from correcting the diet.

(iii) *Ultra-violet Rays*.—Palm, in 1890, drew attention to the relationship between lack of exposure to *sunlight* and the incidence of rickets. It is well known, too, that rickets tends to appear after the dark winter months, and to be healed during the bright summer months.

Rickets can be prevented or cured without improving the deficient diet by exposing the body to the quartz mercury vapour lamp or sunlight. The *ultra-violet rays* are the active ones, chiefly those having a wavelength of 300 $\mu\mu$ or less. The radiation may be applied to any part of the body, and yet can cure the rickety bones generally. The ultra-violet rays act by converting an inactive sterol (*7-dehydrocholesterol*) in the skin, into the active anti-rachitic substance vitamin-D_3. [The activity of vitamins-D_2 and -D_3 is qualitatively and quantitatively the same in man.]

Mode of Action of Vitamin-*D*.—The vitamin promotes the *absorption of calcium and phosphate from the bowel.* When the vitamin is absent adequate absorption does not occur; there is an excessive loss of these substances in the fæces, and the *calcium and/or phosphate level of the serum falls*. The bones are consequently supplied with blood which is deficient in the mineral constituents which are essential for ossification, with the result that soft and irregularly formed bone is laid down. The vitamin also acts in some more *direct manner on the bones*, promoting calcification. Administration of the vitamin causes degeneration of the growing cartilage cells of the epiphysis; this is followed by invasion of the older cartilage zone by blood vessels which bring the necessary calcium salts and thus permit ossification to proceed. In a short time the masses of irregularly arranged cartilage cells of the epiphysis are reduced to the normal short, orderly parallel columns.

There is a seasonal variation in the inorganic phosphate of the serum in children; it falls during the winter, reaching its minimum in March, and rises again to its maximum in the late summer. Greater absorption of phosphate takes place in the summer owing to synthesis of vitamin-*D* in the skin by ultra-violet rays.

The relationship of vitamin-*D* to the *structure of teeth* is discussed on p. 1013.

Tetany may occur as a complication of rickets owing to the lowered serum calcium (cf. p. 1007).

Toxic Effects of Vitamin-*D* Excess.[1]—In recent years large doses of vitamin-*D* have been given in the treatment of a number of diseases, most successfully in lupus vulgaris. The dosage administered to *adults*, *e.g.* 5 mg. (200,000 i.u.) daily for several months may produce toxic effects which closely resemble those of parathormone overdosage. In *children* 1 mg. daily may cause similar effects. The symptoms include anorexia, nausea, vomiting, headache, drowsiness, polyuria, and polydipsia. Both serum Ca and serum P are increased, as is the excretion of these substances in the urine; calcium is deposited in the heart, large blood vessels, lungs, renal tubules, and in other soft tissues; osteoporosis occurs. Signs of renal failure and hypertension may develop. The effects are reversible if vitamin-*D* administration is stopped in time.

Osteomalacia.[2]—This disease of adults is rarely met with in Western countries, but is fairly common in some parts of China, where it affects 1–3 % of child-bearing women. The disease is limited to the female sex, and the symptoms usually appear during pregnancy or the puerperium; sometimes the first manifestations appear at puberty. The bones—especially the pelvic girdle, ribs, and femora—become soft, painful, and deformed. Deformation of the pelvis may be so severe as to necessitate Cæsarean section to procure delivery of the child. Symptoms usually recur with each succeeding pregnancy, but tend to clear up after lactation is completed.

The diet of these women is very inadequate; it contains an excess of cereals, very little fat, and the calcium content of the food is minimal. The serum calcium is low, *e.g.* 6 or 7 mg–%; the inorganic phosphate of the serum is variable—it is low or normal. Administration of calcium salts alone does not relieve the condition; but the administration of cod-liver oil (combined as far as possible with an otherwise adequate diet) cures or prevents the disease and restores the normal serum calcium level.

Osteomalacia thus appears to be a form of adult rickets which is due to inadequate absorption of calcium owing to a *deficiency of vitamin-D and of calcium* in the diet; the onset of symptoms during pregnancy and lactation is presumably due to the demands made by the fœtus (which has " first call " on any available calcium) and by the mammary glands on the calcium reserves in the bones of the mother. There is no evidence of disease of the ovaries or of the parathyroid glands. Owing to the low level of serum calcium, *tetany*, which may be very severe, is a common complication. It is interesting to note that the bones of the fœtus show no signs of rickets, but only rarefaction (*osteoporosis*).

In addition to the form of osteomalacia described above another similar disease due also to calcium and vitamin-*D* lack in the diet may occur in adults of either sex who are living under extreme famine conditions. Many such cases have been reported from European countries after both the first and second world wars. There are also rare syndromes clinically indistinguishable

[1] Anning *et al.*, *Quart. J. Med.*, 1948, *17*, 203.

[2] Maxwell, *Proc. roy. Soc. Med.*, 1930, *23*, 639. Albright *et al.*, *Medicine*, 1946, *25*, 399.

from osteomalacia, which are the result of other metabolic defects and not of calcium or vitamin-*D* deprivation.

DIET AND THE TEETH

Structure of Tooth.—Tooth consists of: (i) *Enamel*, which covers the crown. (ii) *Dentine* or the main substance. (iii) *Cement*, a layer of bone which invests the root. (iv) *Dental periosteum*, which lines the socket and so securely fixes the tooth. (v) *Pulp* (see Fig. 639).

Enamel.—This consists of elongated hexagonal prisms which are set vertically on the surface of the dentine. The prisms are initially fibrous, but later on become completely calcified by impregnation with calcium phosphate and carbonate.

Dentine.—This is a dense bone-like substance (but without lacunæ or Haversian canals) which forms the main substance of the tooth. It contains numerous fine tubules which radiate upwards and outwards from the pulp cavity and branch and become finer as they proceed. The tubules contain processes of the *odontoblasts* (the superficial cells of the pulp). The tissue between the tubules is normally quite homogeneous.

Cement.—This is a layer of lamellated bone which covers the tooth below the point where the enamel stops; it contains lacunæ but no Haversian systems. It is covered superficially by the dental periosteum.

Pulp.—This is a jelly-like tissue containing branched cells, blood vessels and nerve fibres. Covering its surface and adjoining the dentine is a continuous layer of epithelial cells, the *odontoblasts*, which send processes, as already described, into the dentinal tubules.

Relation of Diet to Tooth Structure.[1]—Experiments on dogs suggest that the factors which regulate the calcification of bone are also concerned with the normal development of teeth. When the *vitamin-D* content of the diet is deficient, especially when this is combined with an *excess of cereals* and a *shortage of calcium and phosphorus*, the jaw bones become soft and deformed and all parts of the tooth suffer. The teeth erupt irregularly and their surface is rough instead of being smooth and shiny. The enamel and dentine are thin and poorly calcified; the dentine shows many irregular interglobular spaces (Fig. 640). The diet of the mother during pregnancy also has a profound influence on the subsequent condition of the teeth.

If a tooth is injured, *e.g.* by filing the cusps several times a week for some weeks, it displays considerable powers of repair. When the injury involves the enamel only, the underlying dentine becomes altered and translucent in appearance. When the dentine is injured as well, a response takes place in the related part of the pulp; the odontoblasts proliferate and lay down fresh hard so-called "secondary dentine" which protects the underlying soft tissues. The quality of this secondary dentine depends on the adequacy of the diet which the animal is receiving during the period when it is being laid down.

According to May Mellanby there is in *children* too a close relationship between the structure of the tooth and its liability to decay (*caries*), *i.e.* the

[1] May Mellanby, *M.R.C. Sp. Rep., Diet and Teeth*, pt. i. 1929; pt. ii. 1930; pt. iii. 1934. Mellanby *et al.*, *Brit. Med. J.*, 1952, i, 7.

tooth which initially is badly formed is more liable to suffer in this way. The addition of vitamin-*D* to the diet and the omission of cereals tend to check the progress of dental decay and to promote healing by causing calcification of the carious areas. In 10% of cases, however, caries develops in teeth that are structurally sound. It is suggested that though (in this group) the diet was adequate when the tooth was formed, it was defective when injury was

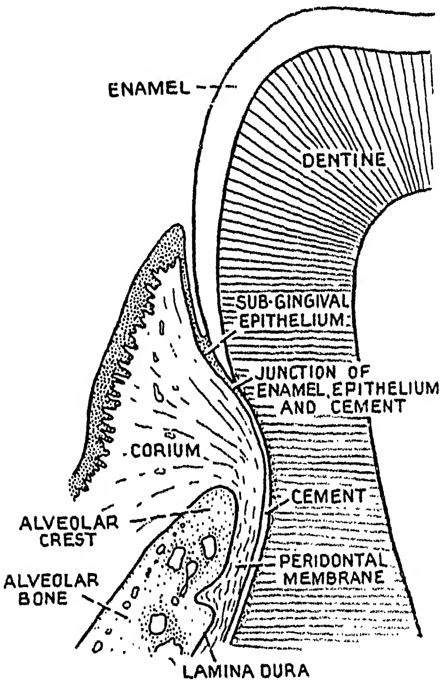


Fig. 639.—Diagram of the Cervical Region of a Tooth and the Related Periodontal Tissues. (After May Mellanby, *Diet and the Teeth*.)

inflicted on it; consequently the normal response did not take place and decay occurred.

The causation of human caries is, however, complex. In the Kangra Valley of the Punjab, for example, the teeth of children with gross rickets and whose mothers had osteomalacia were remarkably good; by contrast in England children with well-formed bones often have many carious teeth.

The mechanisms which may be concerned in the production of clinical caries are critically discussed by Pincus [1] who has shown that changes in the *organic* constituents of the teeth are of major importance, especially those affecting the dentine *protein*.

[1] Pincus, *Brit. med. J.*, 1949, ii, 358.

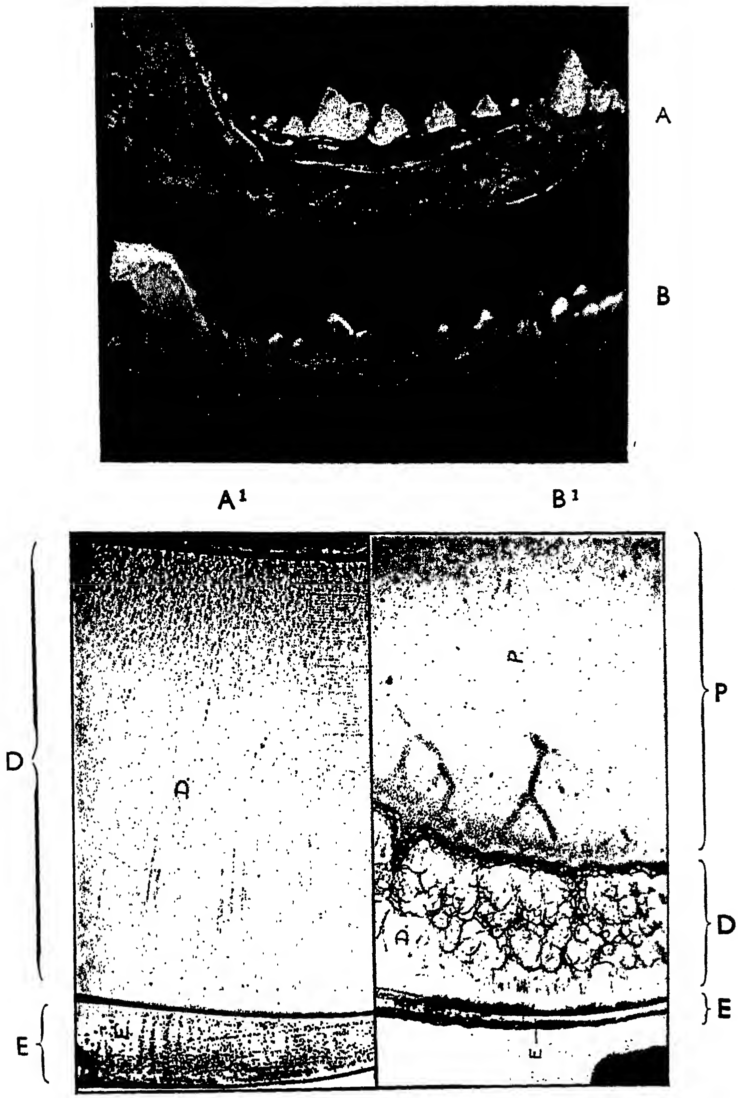


FIG. 640.—Effect of Vitamin-*D* Deficiency on the Structure of Dogs' Teeth. (May Mellanby, *Diet and Teeth.*)

A = Lower jaw of puppy that received cod-liver oil (rich in vitamin-*D*). The teeth have erupted regularly and have a white shiny enamel.
B = Lower jaw of a puppy of the same age that received olive oil (no vitamin-*D*). The enamel is very defective.
A^1 = Microscopic structure of tooth of puppy A. Note the thick well-calcified enamel and dentine.
B^1 = Microscopic structure of tooth of puppy B. The enamel and dentine are thin and badly calcified. The latter shows many interglobular spaces.
In A^1 and B^1, E = enamel, D = dentine, P = pulp.

THE THYMUS[1]

The thymus in man develops from the third branchial clefts of both sides; the two outgrowths come into contact in the middle line in front of the trachea to form a single lobed body.

Structure.—The thymus consists of lobules made up of an outer deeply staining cortex of closely packed lymphocytes, and a central faintly staining medulla which consists of large branched reticulum cells, a few lymphocytes, and the concentric corpuscles of Hassall which consist of a hyaline centre surrounded by layers of flattened cells with poorly staining nuclei.

According to Hammar, the thymus is an ectodermal structure which later becomes invaded by lymphocytes derived from the mesoderm. He regards Hassall's corpuscles as epithelial cells which have undergone a peculiar form of differentiation.

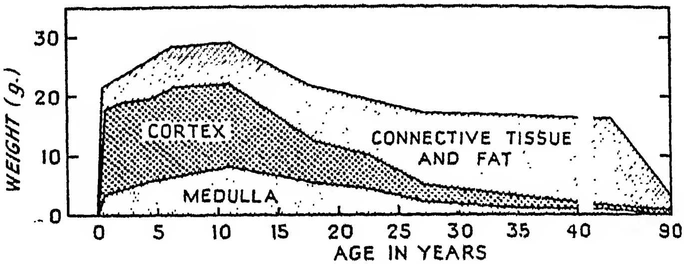


FIG. 641.—Changes in Weight of Thymus at different ages. (Boyd, *Amer. J. Dis. Child.* 1936, *51*, 313.)

Factors Influencing Size of Thymus.—(i) *Relation to Puberty and Gonads.*—The thymus continues to enlarge until the age of *puberty*, and reaches its greatest development about the tenth or the twelfth year; it then steadily becomes smaller and atrophies as age advances (Fig. 641).

After castration in young animals the thymus persists and does not undergo its customary involution. It is suggested that the regression of the thymus at puberty is due to the internal secretions formed by the now mature gonads.

(ii) Injection of pituitary *growth* hormone causes growth of the thymus gland in experimental rats. The gland is commonly enlarged in acromegaly.

(iii) Thymus and lymphoid tissue unite with and inactivate *thyrotrophin* (p. 980). The thymus is commonly enlarged in Graves' disease.

(iv) Injection of *ACTH* or *adrenal corticoids* reduces the lymphocyte content of the thymus and lymph nodes. There is a coincident decrease in the circulating blood lymphocytes.[2] The thymus is enlarged in Addison's disease.

The thymus and the lymphoid tissues generally rapidly atrophy in *starvation* and states of malnutrition. The change is probably due to increased

[1] Park and M'Clure, *Amer. J. Dis. Child.*, 1919, *18*, 317. Hammar, *Endocrin.*, 1921, *5*, 543, 741. Andersen, *Physiol. Rev.*, 1932, *12*, 1. McEachern, *Medicine*, 1943, *22*, 1.

[2] Yoffey, *Biol. Rev.*, 1950, *25*, 314.

corticoid secretion. Thus intact rats starved for 4 days lose 45% of the protein of their lymphoid tissue; adrenalectomized animals, however, show no such loss. It is suggested [1] that adrenal corticoids are secreted in starvation and "mobilize" and break down tissue proteins. The lymphoid tissues are a rich store of labile protein which is readily hydrolyzed to amino-acids to be used for the maintenance of more essential tissues and for neoglucogenesis.

Functions of Thymus.—The only known function of the thymus is the manufacture (and possibly the destruction) of lymphocytes. Its responses are those of lymphoid tissue in general. Extirpation or injection of extracts of this gland in normal animals has produced no significant reproducible results. There is no decisive evidence that the thymus forms an internal secretion.

The gland is commonly enlarged in *myasthenia gravis*; in some cases removal of the gland has resulted in transient or sometimes more lasting clinical improvement. It has been suggested that in this disease the enlarged thymus may be releasing a curare-like substance (p. 519); extracts of the removed glands, however, have no such pharmacological action.

Hyperplasia of Thymus.—There are cases in which the thymus fails to undergo involution at the proper time or there is renewal of growth after involution has been established. In these latter cases it may attain a weight several times that of the normal organ. Such hyperplastic glands are found in infants, usually unassociated with general lymphoid overgrowth; they occur in older individuals in connection with derangements of the organs of internal secretion, *e.g.* Addison's disease, acromegaly, eunuchoidism, in cases of myasthenia gravis, and in status thymo-lymphaticus.

In infants the abnormal thymus may weigh 60 g. at birth, and constitute an obstruction to the respiratory passages or great veins; in older individuals it is doubtful whether mechanical obstruction is often caused by an enlarged thymus. A condition called *thymic asthma* is described, in which attacks of suffocation occur which are said to be relieved by partial removal of the enlarged gland.

STATUS THYMO-LYMPHATICUS.—In many young persons who have died suddenly after some slight injury or infection which seemed insufficient to cause death, general swelling of the lymphoid structures and enlargement of the thymus are often found at autopsy. It has been supposed that the thymic hyperplasia is related in a significant manner to the occurrence of sudden death. But the information available is inadequate for any judgment to be passed. It is claimed [2] that general lymphoid overgrowth is not regularly associated with the thymic enlargement in these cases, and that there is consequently no justification for the name "status thymo-lymphaticus." Others regard the thymic overgrowth as only part of a "general constitutional inferiority" of the affected person. In male subjects there is a tendency to female secondary sex characters, *e.g.* scanty hair on the face, smooth velvety skin, small external genitals, and rounded thighs. In women the thorax and extremities are slender, the genitals hypoplastic, and menstruation is irregular or absent. The musculature may be flabby, the heart is weak, and the large blood vessels are relatively narrow and thin-walled. The failure of this inadequate circulatory mechanism may be

[1] White, *Recent Progress Hormone Research*, 1949, *4*, 153.
[2] Hamer *et al.*, *J. Path. Bact.*, 1931, *34*, 213.

one of the important causes of sudden death in these subjects. There is no proof that the thymus is concerned in the production of the various clinical manifestations described or in causing the sudden death which may occur. It is suggested that incomplete differentiation of the sex glands may be responsible for the failure of the thymus to undergo involution, and may also account for the other features of the condition.

X

NUTRITION

THE VITAMINS[1]

The term vitamin has undergone important changes in meaning since it was first introduced; as a result it carries with it some imprints of all its different meanings. Forty years ago it was believed that the essential constituents of a diet were: protein (in amounts sufficient to maintain nitrogenous equilibrium in the adult and growth in the young), carbohydrate and fat (these three foodstuffs together must be present in sufficient amounts to yield the full calorie requirements), certain minerals (inorganic ions), and water. But later, when a *chemically pure* diet of this kind had been prepared and administered, the animals died; *natural* food, therefore, contains other, non-calorie-providing, but nevertheless essential, constituents for growth, health, and life. In rats, addition of small amounts of milk to a diet which according to the theories then current was adequate (but actually lethal) preserved health and restored growth; the unknown essential factors in milk were called *accessory food factors* by Hopkins. In the meantime other studies showed that certain other theoretically adequate diets, of which polished rice was the main constituent, produced beri-beri; addition of rice "polishings" to the diet cured the disease. The active principle in the rice polishings was called a *vitamine* (*i.e.* an amine essential to life). When it was discovered that few of the "vitamines" were in fact amines, the word was respelt *vitamin*. The term vitamin gradually came to be defined as a substance of unknown chemical composition, probably a complex organic compound, which must be present in the food in minute amounts to enable growth, health, and life to be maintained. The accessory food factors or vitamins were soon divided into (i) *fat-soluble*, *i.e.* those present in fats and soluble in fat-solvents; (ii) *water-soluble*. The fat-soluble were differentiated into vitamins-*A* and -*D*; the water-soluble into -*B* and -*C*. It was soon found that vitamin-*B* was not a single substance but a mixture of several substances; its title was altered to *vitamin-B complex*, and the individual constituents, as they were isolated, were given distinctive names. Rapidly the chemical identity of the vitamins was worked out and many such substances are now known.[2] They are:

$$\left.\begin{array}{l}\text{Vitamin-}A\\ \quad\text{,,}\quad -D\\ \quad\text{,,}\quad -E\\ \quad\text{,,}\quad -K\end{array}\right\}\text{All fat-soluble}$$

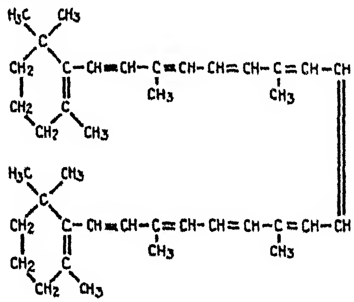
[1] Harris, *Vitamins: A Digest of Current Knowledge*, London, 1951. Bicknell and Prescott, *Vitamins in Medicine*, London, 2nd edn., 1945. Wolbach and Bessey, *Physiol. Rev.*, 1942, *22*, 233.

[2] Most of the vitamins are *groups* of closely related substances, *e.g.* vitamin-*D* group (p. 1009).

Vitamin-*C*
,, -*B* complex, *i.e.* Thiamine
Nicotinic Acid
Riboflavin
Pyridoxin
Biotin
Pantothenic Acid [1]
Folic Acid [1]
Vitamin-B_{12} [1]
and others

The distinctive characteristic of the vitamins is that they are *micro* nstituents of the diet of high biological activity which cannot be replaced by her *normal* dietary constituents. They appear to be indispensable for rtain specific functions of the body; in most cases their mode of action s been fairly clearly defined. They do not contribute to the energy of the dy; in many cases they *mobilize* energy.

β-CAROTENE



VITAMIN A

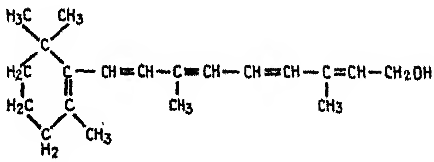


FIG. 642.—Structure of β-Carotene and Vitamin-*A*.

Sometimes an indispensable vitamin for one species (X) can be synthesized by another species (Y); the substance in question is then not a vitamin for species Y. Thus ascorbic acid is an indispensable vitamin in man; the rat, however, can manage without *dietary* ascorbic acid. Both species need and use ascorbic acid in their tissues but while man can only get ascorbic acid by taking it in as part of his food, the rat can synthesize it. Another example: the bacteria in the intestine synthesize several members of the vitamin-*B* group; these vitamins are thus supplied to the body partly by the diet and partly by the intestinal bacteria. If the bacteria are killed off by sulphonamides, signs of vitamin-*B* deficiency may develop because the supplies of dietary vitamin are in themselves inadequate. The generalization

[1] See Emerson and Folkers, *Ann. Rev. Biochem.*, 1951, *20*, 559.

is as follows: indispensable chemical compounds or groupings that can be synthesized in the body (the body includes its cooperative intestinal bacteria) need not be provided as such in the food; indispensable chemical compounds or groupings that cannot be synthesized in the body must be provided as such in the food and constitute the so-called vitamins.

Vitamin-*A*, the -*B* group and -*C* are dealt with below. The other vitamins are discussed elsewhere: -*D*, p. 1008; -*E*, p. 1086; -*K*, p. 151.

Vitamin-*A*.[1]—The vitamin has the formula shown in Fig. 642; a closely related substance is the hydrocarbon *β-carotene* ($C_{40}H_{56}$) which is readily transformed in the cells of the *intestinal wall* into *vitamin-A* ($C_{20}H_{30}O$, an alcohol with a terminal $—CH_2OH$ grouping). Both carotene and vitamin-*A* belong to the general group of *carotenoids*, a series of *yellow-red* pigments characterized by alternate single and double bonds along their main C chain. Vitamin-*A* can now be prepared synthetically on a commercial scale.

As carotene is a precursor of vitamin-*A* it is often referred to as *pro-vitamin-A*.

One international unit (i.u.) of vitamin-*A* is the activity (in animal experiments) of 0·344 μg. of crystalline vitamin-*A* acetate, which corresponds to 0·3 μg. of crystalline vitamin-*A*. One i.u. of carotene is the activity of 0·6 μg. of carotene; this standard recognizes the fact that *dietary carotene has about half the biological activity of vitamin-A*. The explanation of this lower activity may be that vitamin-*A* is absorbed from the intestine more readily than carotene.

SOURCES.—Vitamin-*A* is found as such in the fat of *milk* and therefore in milk products like butter or cream; in *eggs*; in very large amounts in *liver fat*, especially in cod-liver oil and in greatest concentration in halibut-liver oil. The vitamin is absent from purely vegetable fats like linseed oil, olive oil, or coconut oil, and consequently from purely vegetable margarines. All margarine in Great Britain, however, has since the last war been reinforced with added vitamin-*A* (and -*D*). Green vegetables and carrots are free from vitamin-*A* but contain variable amounts of carotene. Both the vitamin and carotene are stable and can withstand the ordinary processes of cooking and boiling.

It is difficult to assign precise values to the vitamin-*A* or carotene content of the foodstuffs. In the case of foodstuffs of animal origin the value varies with the amount of carotene which the animal ingested in its diet; extremely wide variations (up to tenfold) are also found in the vegetable foods. The following average values serve as a guide: milk, 2000 i.u. per pint; butter, 2000 i.u. per ounce; one egg, 200 i.u.; cod-liver oil, 200–13,000 i.u. per teaspoonful; halibut-liver oil, 600–7200 i.u. per drop (20 μg.); carrots, 2000–7000 i.u. (as carotene) per 100 g.; cabbage and green leafy vegetables generally, 1000–2000 i.u. (as carotene) per 100 g.: the white hearts contain less.

REQUIREMENTS.—There is considerable uncertainty about both minimal and optimal requirements of vitamin-*A*. The amounts recommended (*per day* in i.u. of carotene) are: for an adult, 5000; young children and during puberty, adolescence, pregnancy or lactation, up to 6000.[2] A diet containing

[1] Morton, *Ann. Reps. Chem. Soc.* for 1949, *46*, 244.

[2] These values hold when all the vitamin is administered in the form of its precursor, *i.e.* carotene; about half these amounts are needed if the vitamin is given as such.

a half-pint of milk, 1 oz. of butter or vitaminized margarine, and a good portion of green vegetables or carrots daily should be adequate for minimal needs. If the milk ration were increased to one pint daily and cod- or halibut-liver oil were added optimal needs would be easily met.

ABSORPTION.—Vitamin-A and carotene are absorbed from the small intestine together with fat (p. 863); the fat acts as a "carrier" for the fat-soluble vitamin-A and fat-soluble carotene.

Relation of Vitamin-A to Vision.—The phototropism of plants (*i.e.* their tendency to bend towards the light) and the similar phenomenon of free-swimming lowly animal forms moving towards the light depends on the action of the blue end of the light spectrum; these waves are absorbed by a group of yellow pigments, the carotenoids, which constitute the

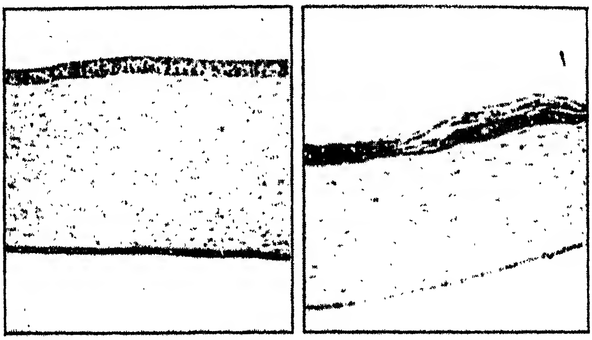


FIG. 643.—Corneal Changes in Vitamin-A Deficiency.
(E. Mellanby, *Nutrition and Disease*, 1934.)

A = Section of normal cornea. B = Cornea from a rabbit fed on a diet deficient in Vitamin-A and carotene. Note thickening and keratinization of the epithelium of the cornea in B.

"chemical eye" of anatomically eyeless plants or invertebrates. In the vertebrates carotenoids are found in the anatomical eye, both in the rods and in the cones.

(i) *Rods.*—The rhodopsin (visual purple) of the rods is a red-coloured carotenoid pigment which is conjugated with protein. Under the influence of light, rhodopsin is split into a colourless protein and the yellowish compound, *retinene*, which is the *aldehyde* of vitamin-A. Subsequently this retinene is converted into vitamin-A, which in turn is combined with protein and reconverted into rhodopsin.

(ii) *Cones.*—Carotenoids have been extracted from the macula. In chickens the pigment is a violet substance called *iodopsin*: from the yellow spot (macula lutea) of man a pigment has been extracted, said to resemble the *xanthophyll* of leaves, which has the empirical formula $C_{40}H_{54}(OH)_2$.

It is clear from this work, that carotenoids, including vitamin-A are related to mammalian vision.

DARK ADAPTATION.—On passing from a brightly illuminated to a dark environment, the speed of *dark-adaptation* (*i.e.* the ability to see better in the dark) depends on the speed of reformation of the broken down visual purple;

such adaptation is impaired in vitamin-*A* deficiency. Dark-adaptation is tested as follows :the eyes are brightly illuminated and the light switched off. One of two measurements is then made : (*a*) the time that elapses before an object of standard brightness is seen ; (β) the degree of brightness necessary for an object to be visible at the end of a fixed time. Though the results are considerably influenced by the *psychological* condition of the subject, depressed dark-adaptation (*i.e.* below normal standards) is the most sensitive index of human vitamin-*A* deficiency.

Results of Vitamin-*A* Deficiency.—ANIMAL EXPERIMENTS.—Vitamin-*A* lack produces striking and characteristic pathological changes in experimental animals. *Young* rats fed on a diet complete in all respects except for the absence of vitamin-*A* cease to grow, lose weight, and finally die. It should be stressed that lack of any one of the vitamins arrests growth in

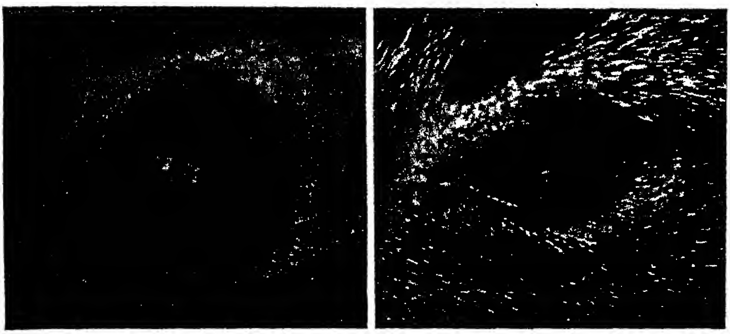
A  B

FIG. 644.—Xerophthalmia following Vitamin-*A* Deficiency. (E. Mellanby, *Nutrition and Disease*, 1934.)

A = Normal eye. B = Affected eye.

young animals. Fully grown animals deprived of vitamin-*A* may survive for months, but finally succumb to some intercurrent *infection*.

Striking pathological changes are found experimentally in the eyes, intestine, and respiratory tract. The outstanding histological feature is the tendency for *stratified epithelia to become greatly thickened* and for *columnar epithelia to be transformed into transitional or stratified epithelia.*

(i) *Eyes.*—The lachrymal glands cease to produce tears. The corneal epithelium becomes thickened, dry and wrinkled (Fig. 643) and secondarily may undergo necrosis or become infected. Inflammatory processes occur in the conjunctiva and may involve the anterior and posterior chambers of the eye, leading to complete blindness (Fig. 644). The eye disorder is referred to as *xerophthalmia*, and is a specific sign of experimental vitamin-*A* deficiency.

(ii) *Alimentary Canal.*—The cells of the salivary glands do not secrete, and appear shrunken ; the epithelium of the ducts is proliferated and the lumen is partially occluded. The mucus-secreting cells of the intestine are atrophied and the tips of the villi are necrosed ; masses of bacteria may be found filling the lumina of the glands.

(iii) The upper *respiratory tract*, particularly the nasal passages, trachea,

and bronchi, shows a transformation of the lining cells into a stratified epithelium of flattened cells which undergo extensive keratinization. [Similar changes occur in the *vagina* (p. 1087) and *gums*.]

(iv) *Resistance to Infection.*—As a result of the structural changes in many epithelia and other tissues their *local resistance to infection* is reduced. From these sites, organisms may pass into the blood stream; bronchopneumonia, enteritis, and inflammation of the eyes may develop and prove fatal. In mice fed on a diet deficient in vitamin-*A*, the introduction of suitable measured doses of mouse-typhoid bacilli may give a mortality rate of 80–100%. When the experiment is repeated on mice fed on the same diet but with the addition of large amounts of vitamin-*A*, the mortality rate is only 10–20%, though in all external appearances the two batches of mice may appear equally normal.

(v) Vitamin-*A* deficiency in dogs leads to marked overgrowth of certain bones, especially the skull and vertebral column. Owing to the resulting compression of the spinal roots and cranial nerves complex nervous symptoms develop. Vitamin-*A* may be a factor regulating normal bone growth.[1]

Studies in Man.[2]—EXPERIMENTAL PURE VITAMIN-*A* DEFICIENCY.—Twenty volunteers were kept for long periods on a diet free from vitamin-*A* or carotene. Many investigations were carried out but the only positive finding was an impairment of *dark adaptation* (p. 1022). The very different results in animals and the clinical findings to be described below may be due in part to species differences, the duration of the deficiency and the presence of other conditioning factors, *e.g.* multiple deficiencies or infections.

CLINICAL FINDINGS.—(i) *Xerophthalmia.*—The associaton of this condition in man with vitamin-*A* deficiency is firmly established. In Denmark, during the first world war, the fats rich in vitamin-*A* were exported, and the populace lived largely on margarine and skimmed milk; several outbreaks of xerophthalmia and bronchopneumonia were reported. Fifteen hundred cases of xerophthalmia occurred in Japanese children who were fed on inadequate diets; the condition was cured by the administration of cod-liver oil.

(ii) *Skin Changes.*—Changes in the *skin* often occur in vitamin-*A* deficiency in man. Thus, in many gaols, asylums, hospitals, and schools in Africa, China, and Ceylon the inmates frequently suffer from dryness of the skin and papular eruptions due to destruction of the ducts of the sweat glands and the hair follicles; the condition is called "toad skin" in Ceylon. It usually precedes the development of xerophthalmia, and like the latter condition it is rapidly cured by means of cod-liver oil.

(iii) *Infections.*—There is no clear evidence that vitamin-*A* lack *in man* lowers resistance to the common infections; excess vitamin-*A* certainly does not raise general immunity above normal.

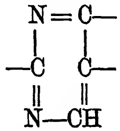
Vitamin-*B* Group.—This group consists of a series of water-soluble organic substances which are found in all the tissues of all species from the bacteria, protozoa or yeasts up to the highest mammalian forms. Most of the members of the group are constituents of fundamental tissue *enzyme systems*, *e.g.* those involved in the oxidation of the foodstuffs, and are therefore indispensable for the normal functioning of all tissues. As might be

[1] E. Mellanby, *Proc. roy. Soc. B.*, 1944, *132*, 28.
[2] *Brit. Med. Res. Council Sp. Rep.*, 1949, Series 264.

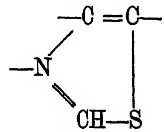
expected, the best studied members of the group, *i.e.* thiamine, riboflavin, and nicotinic acid are generally found together in foodstuffs but not necessarily in the same proportions. Most members of the vitamin-*B* group can be synthesized by the intestinal bacteria (cf. p. 835).

The chemistry, requirements, and functions of each member of the -*B* group are discussed below. The results of vitamin-*B* deficiencies are considered on pp. 1029 *et seq.*

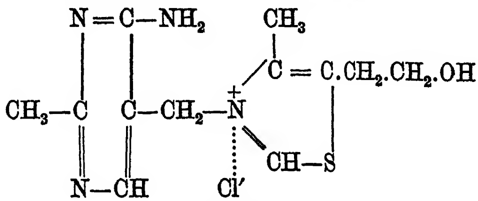
I. **Thiamine** [Aneurin. Vitamin-B_1].



Pyrimidine



Thiazole



Thiamine

FIG. 645.—Structure of Thiamine (Vitamin-B_1).

The molecule (shown in Fig. 645 as the *hydrochloride*) contains a *pyrimidine* and a *thiazole* (sulphur-containing) ring system. The international unit is equivalent to 3 μg. of thiamine.

SOURCES.—The best sources are *cereals*, *pulses*, and *yeast*. In the cereals thiamine is found mainly in the germ and bran. Whole-meal flour prepared from the entire grain, or flour of a high degree of extraction contains much more thiamine than fine white flour (p. 1058); likewise, so-called polished rice (*i.e.* rice from which the husk has been removed) is deficient in thiamine. The vitamin is uniformly distributed in the pulses (peas, beans, lentils). It is present in small amounts in meat, milk, and vegetables. The synthetic vitamin is available commercially.

PROPERTIES.—Thiamine withstands procedures such as drying, pickling or brewing without loss of potency. Cooking processes such as a short period of boiling at 100° C. are not harmful, but rapid destruction takes place in the autoclave at 120° C. Canning in an alkaline medium leads to serious damage to the vitamin owing to its oxidation to *thiochrome*. In the *dried* form the vitamin is very stable and withstands heating at 120° C. for as long as 24 hours.

REQUIREMENTS.—The amount of thiamine needed by the tissues (in the form of *diphospho*thiamine is directly related to *body weight*, the *metabolic rate*, and the level of *physical activity*, and depends on the composition of the *diet*. As thiamine is specifically employed in carbohydrate utilization the amount required is directly related to the amount of glucose which the diet yields. For simplicity of calculation the protein intake is added in with the carbohydrates and the combined calorie value is called the *non-fat calories* (N.F.C.).

'linical studies have shown that when the daily thiamine intake exceeded ·47 mg. per 1000 N.F.C., beri-beri did not occur. The minimum intake, llowing a margin for safety, is estimated at 0·66 mg. per 1000 N.F.C.; the *otal* daily minimum intake on an average diet containing 3000 Cal. of which ver 2000 Cal. are derived from non-fat sources would thus be about 1·5 mg.; he optimal total daily intake might be 2 mg. When the B.M.R. is raised *e.g.* in febrile states) and in people doing hard manual work, a larger intake s needed. It must always be borne in mind, however, that thiamine is synhesized in unknown and variable amounts by the intestinal bacteria. Signs of thiamine lack can be precipitated on what appears to be an adequate diet y administering sulphaguanidine and related sulphonamides which kill off he intestinal bacteria.

RÔLE IN TISSUES.—In the tissues thiamine is found in the phosphorylated orm as *diphosphothiamine*, which is a coenzyme for reactions which involve *lecarboxylation* of acids. Such decarboxylations (and the reverse process alled CO_2 fixation) play an important part in the dissimilation of the netabolites which enter the common metabolic pool. The reactions in which liphosphothiamine is involved are:

(i) Pyruvic acid ⇌ Acetic acid + CO_2 (p. 850).

(ii) Oxaloacetic acid ⇌ Pyruvic acid + CO_2 ("carboxylase" reaction, p. 850).

(iii) Citric acid ⇌ α-Ketoglutaric acid + CO_2 (p. 851).

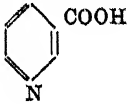
(iv) α-Ketoglutaric acid ⇌ Succinic acid + CO_2 (p. 852, Fig. 559).

All these reactions are stages in *carbohydrate utilization*. It will be recalled hat the brain and probably all nervous tissues use blood glucose as their rimary source of energy. It is not surprising to find that carbohydrate netabolism (especially in the brain) is deranged in thiamine deficiency (p. 1030).

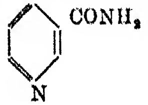
In thiamine-deficient pigeons pyruvic acid accumulates in the brain, specially in the brain stem, because it cannot be disposed of; the concentraion of pyruvic acid also increases in the blood and cerebrospinal fluid. If hiamine is added to a brain slice from such a bird the pyruvate rapidly lisappears.

2. Nicotinic Acid Amide [Nicotinamide].

Nicotinic acid.



Nicotinic acid amide.



SOURCES AND REQUIREMENTS.—The richest sources are liver (15 mg–%), kidney, and yeast; meat is a valuable source containing 5 mg–%: the other important sources are whole meal flour and green vegetables; milk contains only 0.08 mg–%. Of the cereals, *maize* contains least nicotinic acid. The daily need is thought to be about 10 mg. It should be emphasized that nicotinic acid is synthesized by intestinal bacteria. Nicotinic acid is unaffected by cooking.

RÔLE IN TISSUES.—Nicotinamide (which is a pyridine derivative) is found in the tissues as a nucleotide, diphosphopyridine nucleotide (DPN), usually called *coenzyme-I* (Co-I), formed by the combination of adenine

ribose, phosphate, and nicotinamide (Fig. 552, p. 841). There is also a corresponding *tri*phosphopyridine nucleotide (TPN), *coenzyme-II* (Co-II).

Coenzymes-I and -II are the coenzymes for *oxidative enzymes* (*dehydrogenases*); the coenzymes act as intermediate carriers for the hydrogen (2H) released from various substrates by the dehydrogenase enzymes (p. 840).

Co-I is the coenzyme for the following reactions, among others:

(i) Lactic acid→Pyruvic acid+2H (p. 849).

(ii) Pyruvic acid+H_2O→Acetic acid+CO_2+2H (both an oxidation and decarboxylation, p. 850).

(iii) β-Hydroxybutyric acid→Acetoacetic acid+2H (p. 869).

Co-II is the coenzyme for the following reactions, among others

(iv) Citric acid→β-Ketoglutaric acid+CO_2+2H.

(v) Glutamic acid+H_2O→β-Ketoglutaric acid+NH_3+2H (both an oxidation and deamination, p. 885).

All the reactions are shown as a loss of 2H from the substrate on the left, the 2H being passed on to the coenzyme, which is thus converted to the reduced form:

Co-I+2H⟶Co-I.2H.

The same enzyme and coenzyme systems are required by the reactions in reverse, *i.e.* when reduction of the substrate is taking place and the reduced coenzyme is being reoxidized, *e.g.*:

Co-I.2H⟶Co-I+2H

2H+Pyruvic acid⟶Lactic acid.

The nicotinamide-containing coenzymes are thus concerned with many of the important energy-producing reactions of metabolism. Nicotinamide (or nicotinic acid) deficiency leads to metabolic disturbances in many tissues; as might be anticipated, the nervous system is gravely involved. *Pellagra*, the main clinical syndrome of nicotinamide deficiency, is discussed on p. 1037.

3. **Riboflavin.**—Sources and Requirements.—The international unit is equivalent to 3μg. of riboflavin. The main sources are again meat, milk, and wholemeal flour. Clinical studies show that signs of riboflavin deficiency occur when the daily riboflavin intake is 0·2–0·3 mg. per 1000 total Cal.; an intake of 0·35–0·5 mg. per 1000 total Cal. is adequate. On an average diet, a *total* riboflavin intake of 2 mg. daily is probably adequate; more is needed when the metabolic rate is increased. Riboflavin is synthesized by intestinal bacteria.

Structure.—Riboflavin is a yellow-green pigment consisting of a three-ring system (*iso-alloxazine*) combined with an alcohol derived from ribose (*ribitol*) (Fig. 646).

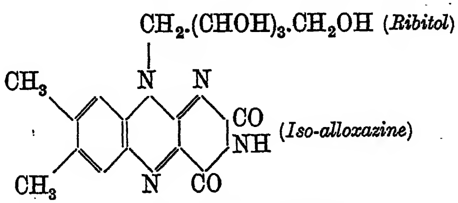


Fig. 646.—Structure of Riboflavin.

Rôle in Tissues.—In the tissues riboflavin is found as a dinucleotide made up as shown on p. 1028.

Iso-alloxazine   Adenine
| |
Ribitol—Phosphate—Phosphate—Ribose.

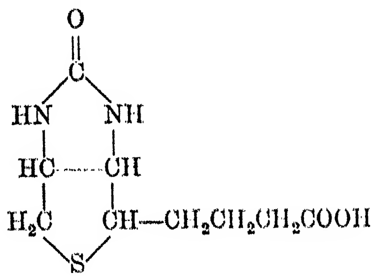
This substance is called flavin-adenine-dinucleotide (FADN), or, for brevity, *flavin*.

Flavin is the prosthetic group of a number of *dehydrogenase enzymes* (the *flavoproteins*) (p. 854). These enzymes catalyse the release of 2H from various substrates, and their flavin group acts as a temporary *hydrogen-carrier.* The hydrogen is passed from the flavin to the cytochromes (p. 854), and thence finally to reactions with molecular oxygen to give water. In particular, those oxidation reactions requiring coenzymes-I or -II as initial hydrogen-carriers (pp. 840, 854) require flavoproteins as the next hydrogen-carrier of the chain leading to molecular oxygen.

As flavin occupies a key position in reactions leading to the oxidation of hydrogen to water, one would expect that dietary deficiency of riboflavin would lead to dramatic disturbances of function. The actual changes observed clinically are described on p. 1034; they are far less striking than might have been anticipated from the metabolic findings recorded above.

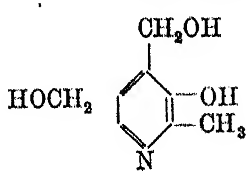
Certain other members of the vitamin-*B* group should be noted, though their rôle in man is still obscure.

4. Biotin.[1]—Its formula is shown below.



Biotin (or a derivative) participates in many CO_2 fixation reactions, particularly the carboxylase reaction leading to aspartic acid.

5. Pyridoxine.—Its formula is shown below.



Pyridoxine and its derivatives (the vitamin-B_6 group) act as coenzymes for transamination reactions (p. 885).

Nothing is known about the effects of pyridoxine insufficiency in man.

6. Pantothenic Acid.—It is concerned (as a dinucleotide, referred to as Coenzyme A) with acetylation reactions (p. 874).

7. Folic Acid and Vitamin-B_{12}.—See p. 198.

[1] Vigneaud, *Science*, 1942, *96*, 455.

Results of Vitamin-*B* Deficiency.—As the members of the vitamin-*B* group are generally found together in foods, *isolated* deficiency of any *single* member of the group rarely occurs clinically, though it has been produced experimentally. In the common clinical condition of *multiple* deficiency the signs of some *one* single deficiency may predominate and mask the results of the other deficiencies. If the outstanding deficiency is made good the latent deficiencies may become manifest in the appearance of new symptoms. It is necessary, therefore, to establish first the clinical picture of single deficiencies before considering the naturally occurring clinical syndromes which probably all involve multiple -*B* deficiencies.

Thiamine Deficiency in Man.—VERY SEVERE DEFICIENCY.[1]—The effects of pure thiamine deficiency have been studied in a few women. The extremely unattractive diet that had to be employed consisted of white flour, sugar, tapioca, starch, cheese, butter, hydrogenated fat, tea and cocoa, with added vitamin-*C*, halibut-liver oil, iron, calcium and autoclaved brewer's yeast (to supply other members of the -*B* group); the daily thiamine intake was 0·15 mg. After *two weeks* the first symptoms appeared, consisting of fatigue and loss of appetite; during the *second month* there was decreased activity, apathy, nausea, decreased food intake, and a fall of blood pressure. During the *third month* the consumption of food was further decreased (the calorie intake was then under 1500 Cal.) and the body weight fell slightly. There was dizziness, the heart rate was slow at rest but excessively rapid on exertion, the voltages of the electrocardiogram, especially the T wave, were decreased; uncontrollable *vomiting* occurred; the motility of the stomach and large intestine was diminished. There was a greatly decreased capacity for work, *weakness* in the legs, and *decreased reflexes*. Some of the subjects became very depressed and showed other *mental changes*; they complained of soreness of the muscles and *numbness* of the legs, muscular flaccidity, precordial distress, and dyspnœa.

The clinical picture of acute thiamine deficiency can be made clearer by giving details of the symptoms in one active, vigorous, and industrious woman subject.

At 30 *days*: no appetite; weakness.

At 70 *days*: nausea constantly present; weakness severe. She only carried out light duties and often did not complete her tasks. She was apathetic and confused, and at times complained of *numbness* and *tingling* in the legs.

At 90 *days*: spontaneous exertion almost ceased. *Tactile sensibility was impaired*; the muscles of the calves were tender; she could not get up readily from the squatting position.

At 110–120 *days*: no appetite, *nausea, vomiting*; *apathy, vagueness and confusion* were conspicuous. B.P., 110/65. Heart rate, 55–65 at rest, 90–120 on slight exertion. Impaired sensation of pain, tactile localization and discrimination over the legs; *vibration sense impaired* at the ankles. All the leg movements were weak and the quadriceps extensors were almost completely *paralysed*. The *knee-jerk* and the *ankle-jerk were absent*.

It will be noted that signs of *organic* changes in the nervous system (anæsthesia, loss of reflexes, and paralysis in legs) made their appearance late; they resemble the *neuritic* changes found in beri-beri. Attention must

[1] Williams *et al.*, *Arch. int. Med.*, 1940, *66*, 785; 1942, *69*, 721; 1943, *71*, 38.

be drawn to the absence of œdema, cardiac dilatation, or changes in the skin and tongue. As might be expected, the rapidity of development of the symptoms varied to some extent with the degree of muscular activity. The terminal state, especially the nausea, vomiting, and mental changes, resembles the findings in *cerebral beri-beri.*

The symptoms described can probably be attributed solely to thiamine deficiency. The addition of thiamine, without any other change in the diet, produced improvement which was sometimes sudden, but commonly (if organic changes had developed) very gradual. Thus, a single injection of 1 mg. of thiamine rapidly relieved some of the symptoms: the nausea and vomiting ceased; the diet which had previously appeared revolting was eaten freely; and apathy disappeared. With sustained intensive thiamine therapy (*e.g.* 15–60 mg. of thiamine daily) strength returned, the mental state brightened, and the signs of organic nerve damage very slowly cleared up. If the subject was put on a thiamine intake of 0·95 mg. daily, recovery

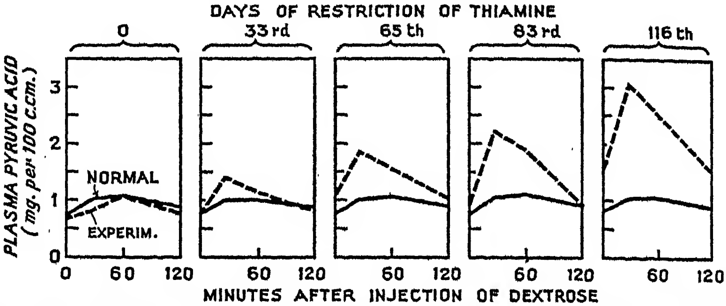


FIG. 647.—Metabolic Changes in Acute Severe Thiamine Deficiency. (Williams *et al., Arch. int. Med.*, 1943, *71*, 46.)

Changes in blood pyruvic acid after intravenous injection of test dose of glucose.
Continuous line: normal response. Dotted line: response after 0, 33, 65, 83, 116 days of thiamine deficiency.

was partial; on a daily intake of 2 mg. full recovery finally occurred and sometimes unusual vigour was displayed. These last data are a guide to minimal and optimal thiamine requirements in man.

BIOCHEMICAL CHANGES.—On a *normal* thiamine intake, the blood thiamine level is 3–9 μg–%; one-third of any injected thiamine is excreted in the urine. After ingestion or intravenous injection of glucose there is only a very slight rise in the level of blood pyruvic acid (normal 0·5–1·0 mg–%) (Fig. 647). In *severe thiamine deficiency, e.g.* on an intake of 0·1 mg. per 1000 Cal., the blood thiamine level fell to under 3 μg–%; the excretion of thiamine in the urine progressively declined to zero after the 10th day. As previously mentioned (p. 1026), in thiamine deficiency carbohydrate metabolism in part *stops short* at the pyruvic and lactic acid stages; if glucose (0·4 g. per kg.) is injected intravenously the blood pyruvic level rises temporarily to an abnormally high level (up to 2–3 mg–%) (Fig. 647), as does the blood lactic acid.

MODERATE THIAMINE DEFICIENCY.—The effects of a thiamine intake of 0·45 mg. daily (less than 0·2 mg. per 1000 Cal.) were studied in 11 normal subjects for periods up to 131 days. For the first 56–84 days the symptoms consisted mainly of *changes in the mental state* and *subjective nervous manifestations.* The subjects became irritable, depressed, quarrelsome, or fearful

that some misfortune was about to befall them; two subjects felt that life was no longer worth living and threatened to commit suicide. Work was inefficiently done; they carried out instructions in a dull and confused way, dropped their needles during sewing and broke the crockery while washing up. They were full of complaints of headache, backache, sleeplessness, painful menstrual periods and a frequent feeling of *numbness* in the hands and feet. The appetite was poor, and nausea, *vomiting*, and epigastric distress were present from time to time. Attention must again be drawn to the resemblance between the changes just described and those of mild cases of cerebral beri-beri (p. 1036). Gastro-intestinal motility was radiographically normal. There was some *anœmia* (*e.g.* red cell count, 3–3·5 million) of a macrocytic hyperchromic type, accompanied by hyperplasia of the red bone marrow. The *plasma protein* concentration was lowered (*e.g.* to 5%); this fall was sometimes associated with a slight pitting *œdema*. Body weight did not decrease.

Recovery occurred *very slowly* with intensive thiamine therapy.

Clinical Vitamin-*B* Complex Deficiencies.—The principal clinical syndromes resulting from lack of the vitamin-*B* complex are summarized below:

Beri-beri: due primarily to thiamine lack. Three main clinical types are recognized: (i) neuritic (dry); (ii) cardiac (wet); (iii) cerebral (Wernicke's encephalopathy).

Pellagra: due mainly to lack of nicotinic acid; a high maize diet is often an aggravating factor.

Ariboflavinosis: the chief changes attributed to riboflavin lack are angular stomatitis, glossitis, scrotal dermatitis, keratitis, defective vision from retrobulbar neuritis, and painful feet.

Mixed Syndromes of -B_2 Avitaminosis: the term vitamin-B_2 is often used to refer to the members of the vitamin-*B* complex other than thiamine (=vitamin-B_1). The varied clinical manifestations are due mainly to lack of nicotinic acid and of riboflavin.

It is interesting to note that when a large population is exposed to a multiple avitaminosis some develop beri-beri, some pellagra, some ariboflavinosis, and some develop peculiar disturbances not falling exactly into any of these categories. As will be explained beri-beri and pellagra may involve many organs or systems; in individual cases the disturbances may affect predominantly one or other organ. The reason for these wide individual variations in response to a common dietetic deficiency is quite unknown. It should also be emphasized that clinically vitamin-*B* deficient diets are commonly unsatisfactory in other respects; they may be lacking in protein, salts, or other vitamins also.

The Singapore Captivity.[1] Much was learnt about the effects of Vitamin-*B* lack from a study of the tragic experiences of some 35,000 British prisoners of war in Singapore who within a few days of 15th February 1942, changed suddenly from normal adequate British rations to a grossly inadequate diet consisting mainly of polished rice. The Japanese daily scale of rations for prisoners was in g.: polished rice 500, flour 50, sugar 5, cooking fat, 5, meat or fish 50, fresh vegetable 100, canned milk 15, salt 10, tea. The full ration was frequently not issued. The calorie value was 2100–2500. For the first 3

[1] Smith and Woodruff, *Deficiency Diseases in Japanese Prison Camps*, *Brit. med. Res. Council Sp. Rep.*, No. 276, 1951.

years of the 3½ years' captivity the diet was not seriously short of protein, fat, Ca, P, vitamin-*A* and -*C*; but in the last 6 months there was semi-starvation as the total calorie intake fell to 1500–2000. For long periods, however, the diet was deficient in vitamin-*B* complex, certainly in thiamine, riboflavin, and nicotinic acid. The dietary deficiency was aggravated by epidemics of dysentery which probably led to additional loss or destruction of the *B*-vitamins in the bowel.

The main results of this study are summarized in Figs. 648–651. Fig. 648

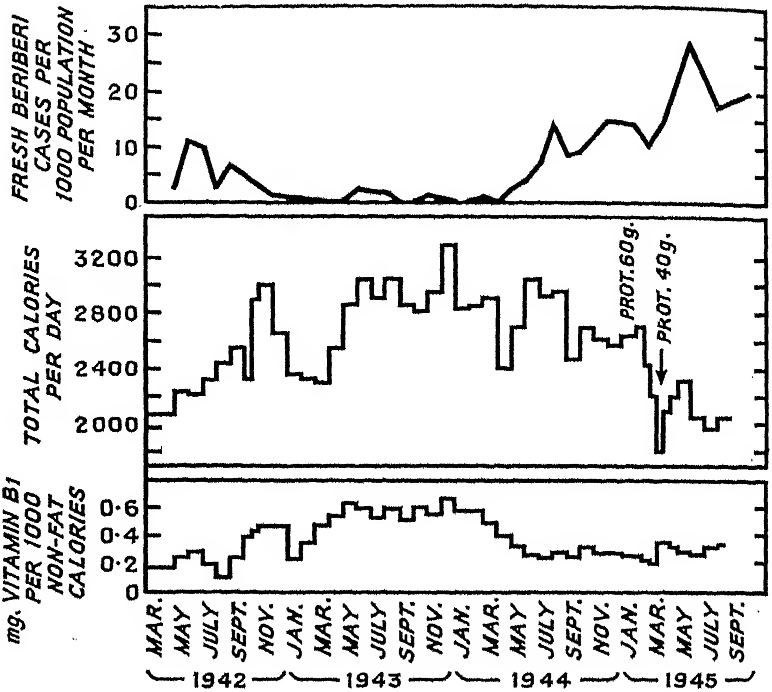


FIG. 648.—Incidence of Beri-Beri in Relation to Total Calorie intake and to intake of Vitamin-B_1 (Thiamine) in mg. per 1000 Non-Fat Calories among British Prisoners of War at Singapore throughout Period of Captivity. (Burgess, *Lancet*, 1946, ii, 412.)

shows the relationship between the incidence of beri-beri and the thiamine intake (in mg. per 1000 N.F.C.) throughout the period of the captivity, *i.e.* from March 1942 to September 1945. During the first seven months, the thiamine intake was 0·2–0·3 mg. per 1000 N.F.C., and numerous cases of beri-beri (neuritic, cardiac, cerebral and mixed) occurred (Fig. 649). The first cases made their appearance after a latent period of six weeks. Between November 1942 and May 1944 the thiamine intake was higher (0·4–0·6 mg. per 1000 N.F.C.) and cases of beri-beri were rare. Later in 1944, as the thiamine intake fell again to 0·2–0·3 mg. per 1000 N.F.C., cases of beri-beri became increasingly numerous. Œdema was now the commonest sign, often

unaccompanied by neuritic changes. Up to this time prevention and cure were obtained by administering thiamine or yeast extract (which contains the whole -*B* complex). A sudden large intake of carbohydrate (*e.g.* polished rice) in prisoners who had been previously starved often precipitated the onset of signs of beri-beri, because the low thiamine intake was inadequate for the large carbohydrate utilization.

After March 1945 the feeding arrangements failed and the prisoners

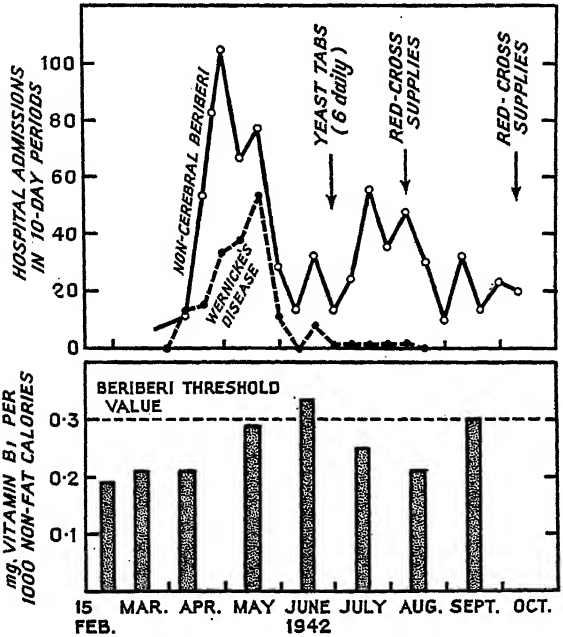


FIG. 649.—More detailed Analysis of Incidence of Beri-Beri (non-cerebral and cerebral) in Relation to intake of Vitamin-B_1 (Thiamine) in mg. per 1000 Non-Fat Calories among British Prisoners of War at Singapore during first Outbreak (Wardener and Lennox, *Lancet*, 1947, i, 14).

The threshold value of 0·3 mg. of thiamine per 1000 non-fat Calories is probably set too low. Cases occurred when the thiamine intake was 0·36 mg.; no cases occurred when the thiamine intake exceeded 0·47 mg. per 1000 N.F.C.

suffered from partial starvation; the total calorie intake was 2000 Cal., the protein intake fell to 40 g. and the dominant finding was famine œdema [1]

[1] There is an interesting reference to famine œdema in Macaulay's essay on Southey's *Colloquies*. He quotes Magendie (from *J. de Physiologie experimentale*) on "a point of physiology connected with the great distress in France in 1817" to the effect that the inhabitants of certain departments "were reduced first to oatmeal and potatoes and at last to nettles, beanstalks and other kinds of herbage fit only for cattle; that when the next harvest enabled them to eat barley bread many of them died from intemperate indulgence in what they thought an exquisite repast; and that a dropsy of a peculiar description was produced by the hard fare of that year." Famine œdema is fully discussed in Keys *et al.*, *Biology of Human Starvation*, Minneapolis, 1950, vol. ii, 921.

which appeared in emaciated patients and was resistant to vitamin-B treatment.

Figs. 650 and 651 show the incidence of the syndromes which were attributed to *vitamin-B_2 lack*. The -B_2 syndromes showed two peak periods

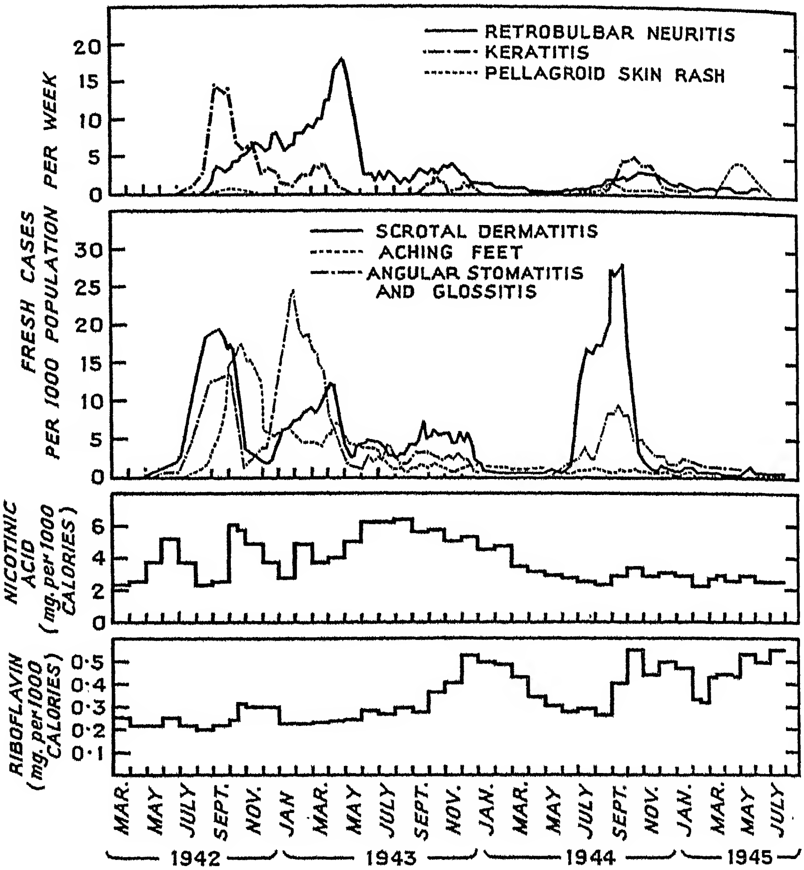


FIG. 650.—Relation of intake of Nicotinic Acid and Riboflavin to incidence of Vitamin-B_2 Deficiency Syndromes among Prisoners of War at Singapore. (Burgess, *Lancet*, 1946, ii, 413.)

Intake of Nicotinic Acid and Riboflavin is expressed as mg. per 1000 total Calories in diet.

of incidence. The first occurred early in the captivity; it appeared after a latent period of two and a half months and lasted longer than the associated beri-beri outbreak, persisting until mid-1943. The second peak occurred in late 1944 and its time course did not coincide with the second beri-beri outbreak. Fig. 650 shows that the various -B_2 syndromes did not follow the same time course either. The "hyporiboflavinosis triad," *i.e.* angular

stomatitis, glossitis and scrotal dermatitis, appeared first, followed by aching (painful) feet, keratitis and paraplegia, and still later by retrobulbar neuritis. A second outbreak of the " triad " occurred in 1944. In the main the incidence of these signs coincided with a low riboflavin intake (0·2–0·3 mg. per 1000 total Cal.); the nicotinic acid intake was adequate. The aching feet are attributed to a complex -B_2 deficiency.[1] The incidence of pellagroid skin

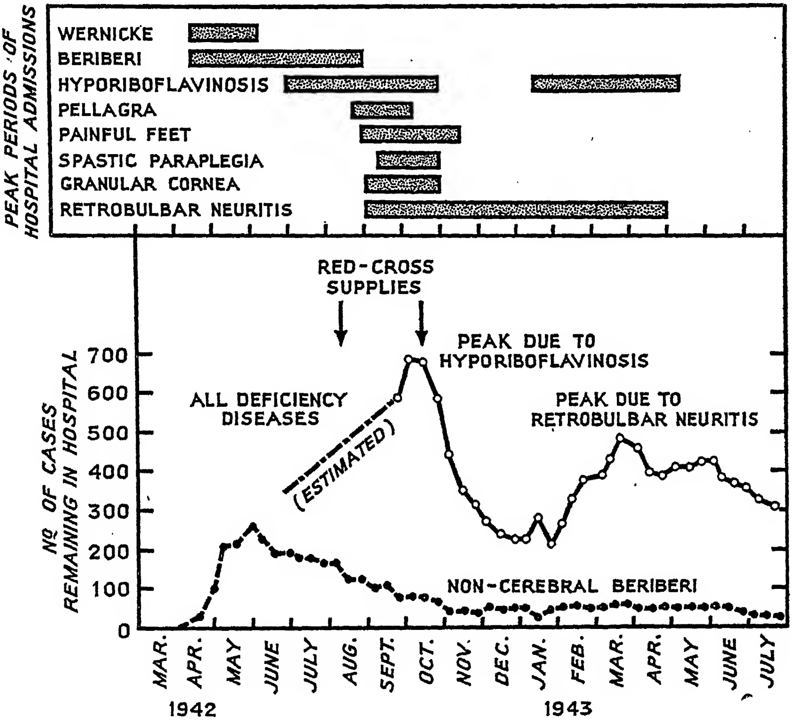


FIG. 651.—Time Relationship of Peak Incidence of Various Vitamin-*B* Deficiency Syndromes among Prisoners of War at Singapore during first Outbreak, March 1942–July 1943. (Wardener and Lennox, *Lancet*, 1947, i, 15.)

" Hyporiboflavinosis " = angular stomatitis, glossitis and scrotal dermatitis. Painful (aching) feet, granular cornea (keratitis), and retrobulbar neuritis are also attributed wholly or partly to riboflavin deficiency.

rash also showed two peaks which coincided with a low nicotinic acid intake.

The clinical findings in beri-beri and pellagra are summarized below.

Beri-Beri.—(1) NEURITIC (DRY) FORM.—The spinal cord and peripheral nerves are chiefly involved. Lesions are present in the ventral horn cells, dorsal root ganglia and the peripheral nerves. There is tenderness of the skin and the deep structures (muscles, bones), wasting and paralysis of muscles and diminution or loss of deep reflexes; the muscles of respiration may be affected also.

[1] Cruikshank, *Lancet*, 1946, ii, 369.

Pellagra.[1]—This disease is endemic in the southern states of the U.S.A., Italy, Rumania and other countries. The findings are summarized below.

(1) GASTRO-INTESTINAL CHANGES.—*Glossitis* appears early in the disease; the tip and lateral margins of the tongue are red and swollen, and later penetrating ulcers develop. Similar changes take place in the mouth, gums and pharynx, and follow the same course. There is complaint of a burning sensation in the mouth, œsophagus and stomach; mucosal changes have been seen in the stomach with the gastroscope. Nausea, vomiting and diarrhœa are frequent. The urethral and vaginal mucous membranes may show changes similar to those present in the mouth.

(2) SKIN lesions may develop anywhere, usually on the dorsum of the hands and feet, axillæ, elbows, wrists, knees, beneath the breasts, and the perineum. The skin is first red and itchy; later it becomes swollen and tense and vesicles develop; finally desquamation takes place, the underlying skin remaining abnormally thickened and pigmented.

(3) NERVOUS CHANGES.—Various forms of *mental* disorders occur, together with "polyneuropathy," *i.e.* symmetrical bilateral involvement of the nerve supply of the lower limbs.

The following facts indicate that pellagra is a deficiency disease:

(i) The diet used in pellagrous households generally consists largely of *maize*, and is poor in meat. Eleven volunteers in a Mississippi prison were given a typical diet consisting of maize meal, white wheat flour, potatoes, salt pork, and syrup; at the end of 6 months 6 had developed pellagra.

(ii) It has been suggested that pellagra is due to *insufficiency* of protein or essential amino-acids. It is found, however, that large amounts of casein —a protein of high biological value—though retarding the onset of symptoms, do not prevent recurrences; on the other hand, the administration of 15 g. of dried *yeast* daily (=7 g. of protein) is a very potent remedy and is equivalent in curative value to 200 g. of lean meat (=45 g. of protein).

(iii) Human pellagra can be controlled—both prevented and cured—by modifications of the diet, *e.g.* by a generous allowance of lean meat or yeast extract; so long as the diet is adhered to strictly, remissions (which otherwise are of frequent occurrence) do not take place.

(iv) Remarkable improvement is obtained by administration of *nicotinic acid*, or *nicotinic acid amide* or certain related substances (*e.g. quinolinic acid*) by any route. In 24–72 hours the redness and swelling of the tongue and mouth and the alimentary symptoms subside. The fiery skin lesions blanch in about 48 hours. Acute mental symptoms disappear; a patient who was previously maniacal may become calm, and confused patients clear. The polyneuropathy, however, is unrelieved by nicotinic acid but is benefited by thiamine. The rôle of maize in the production of pellagra is not understood; it seems to aggravate in some way the harmful effects of nicotinic acid deficiency. Maize may be harmful because it is deficient in tryptophan or because it contains an antagonist to nicotinic acid.

Among the British prisoners of war at Singapore, whose diet was deficient in the whole vitamin-*B* group, the main incidence of pellagra occurred some months after the peak incidence of beri-beri (Fig. 651); the mucous membrane and skin lesions of pellagra appeared first and the changes in the nervous system developed later.

[1] Elvehjem, *Physiol. Rev.*, 1940, *20*, 249.

It seems, therefore, that the major specific signs of pellagra (gastro-intestinal, skin, and mental changes) are due to deficiency of nicotinic acid. In pellagrous areas many cases of psychosis (without other signs of pellagra) are the result of nicotinic acid deficiency and are rapidly relieved by this substance. In most clinical cases of pellagra the diet is deficient in many respects (calories, protein, calcium, iron, vitamin-*A* and the whole -*B* complex). Specific therapy is therefore not a substitute for a general correction of the diet.

Attention has been repeatedly drawn to the synthesis of the vitamin-*B* group by intestinal bacteria. An interesting case has been recorded of an ill-nourished woman who developed dysentery which was treated with sulphaguanidine (p. 1020). A pellagrinous rash appeared in three days which cleared up (in spite of continued administration of sulphaguanidine) by means of injections of nicotinamide.

Nicotinic acid has been employed in the treatment of other forms of psychosis, especially those occurring in old people; beneficial results are claimed. It should be remembered, however, that nicotinic acid increases the cerebral blood flow (*e.g.* by 20%) and may be helpful in this non-specific manner.

Biotin Deficiency.[1]—(1) In *animals* marked symptoms can be produced by a diet rich in egg white; the mechanism concerned is the inactivation of the biotin in the animal's body by the *avidin* of the egg white. In monkeys, biotin deficiency causes thinning of the fur and loss of hair colour; later a heavy scaly dermatitis develops which covers the whole body, but is most marked on the face, arms, and lips. The condition clears up when 20 μg. of biotin are given daily.

(2) A number of *volunteers* were fed on a diet with a minimal biotin content to which was added 200 g. of egg white daily. Fine *scaling of the skin* appeared after four weeks and cleared up about a week later. In the eighth week there was a marked greyish pallor of the skin and mucous membranes (which was out of proportion to the slight anæmia which developed), atrophy of the lingual papillæ and return of the dryness and desquamation of the skin. Symptoms like those of thiamine deficiency (p. 1029), progressively developed from the fifth week, *i.e.* depression, lassitude, muscle pains, hyperæsthesia, loss of appetite, and nausea. Serum cholesterol was raised. Biotin excretion in the urine was lowered from the normal (*e.g.* 25–50 μg.) to 3–7 μg. daily. Ingestion of 75–300 μg. of biotin daily produced relief in a few days with restoration of the appetite, return of the colour of the skin and mucous membranes to normal, and increased biotin excretion in the urine.

Vitamin-*C* (Ascorbic Acid, Cevitamic Acid).—Sources.—The best sources are: (i) *fresh fruits* (in decreasing order): black currant, strawberry, orange, lemon, grapefruit, gooseberry; there is little in pear, plum, grape or most varieties of apple; (ii) *fresh vegetables* (in decreasing order): brussels sprouts, cauliflower, cabbage, tomato, new potato. Vitamin-*C* is present in *small* amounts in fresh milk (more especially human milk), and in raw meat juice; *pasteurized* milk is nearly devoid of the vitamin. The vitamin-*C* content of fresh milk varies with the diet of the cow, being large in amount in summer, when the animal is fed on greenstuffs, and much less in winter. It is absent from dried seeds, but it is formed during the process of *germination*.

[1] Sydenstricker *et al.*, *Science*, 1942, *95*, 176.

Distribution in Animal Tissues.—Vitamin-*C* can be estimated quantitatively in animal tissues by a method which depends on the fact that it decolorizes the dye 2 : 6-dichlorophenol-indophenol. Using this method the vitamin can be demonstrated in the adrenal cortex (p. 944), the wall of the gut, the lens, and the aqueous and vitreous humours. The amount excreted in the *urine* can also be readily determined.

Properties.—It is rapidly destroyed by heating at 100° C in the presence of *oxygen*, especially in an alkaline medium, and by drying. It is therefore absent as a rule from dried, canned or preserved foods unless the process is carried out *anærobically*. During the cooking of vegetables the loss of -*C* can be minimized by *avoiding* too much preliminary shredding, too much cooking water, prolonged cooking, and keeping the vegetables warm for a long time or reheating.

Chemistry.—Vitamin-*C* has been isolated, and is synthesized on an industrial scale for therapeutic use; it is named *ascorbic acid* or *cevitamic acid*. The formula is shown in Fig. 652.

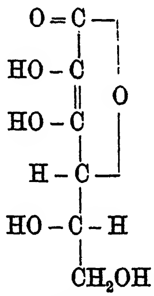


Fig. 652.—Structure of Vitamin-*C*.

Effects of Vitamin-*C* Lack in Man. Experimental Scurvy.—The account which follows is based mainly on the pioneer experiment of Crandon [1] on himself and on a careful study of 20 British volunteers.[2] The latter were maintained for a control period of 6 weeks on a basal diet completely devoid of vitamin-*C* (it contained less than 1 mg. daily) to which 70 mg. of -*C* were added daily. The subjects were then divided into three groups: (i) on the basal diet only (10); (ii) basal diet +10 mg. of -*C* daily (7); (iii) basal diet +70 mg. of -*C* daily (3). Groups (ii) and (iii) remained perfectly fit and well. The detailed findings in the completely deprived British group and in Crandon are considered below.

Blood Changes (Fig. 653).—The average normal vitamin-*C* level in the *plasma* is 0·55 mg./100 c.c.; in the white blood corpuscles the average normal value is much higher, *i.e.* 16 mg./100 g.

(i) On a vitamin-*C* intake of 70 mg. daily the blood levels are unaffected.

(ii) On the vitamin-free diet the vitamin-*C* content fell in the plasma to a negligibly low level, *i.e.* under 0·05 mg./100 c.c. after 5 weeks, and in the white corpuscles to under 1 mg./100 g. after 16 weeks.

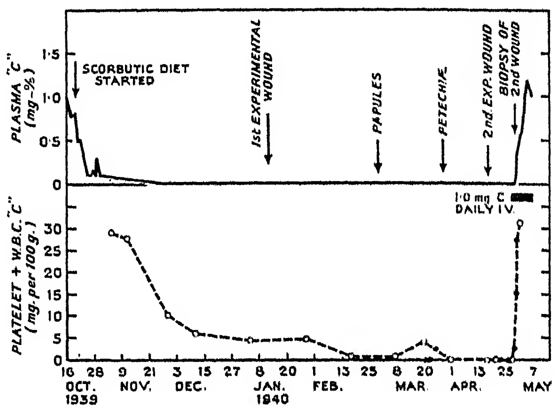
(iii) On a 10 mg. -*C* intake the plasma changes were the same as in (ii)

[1] Crandon, Lund, and Dill, *New Engl. Med. J.*, 1940, *223*, 359.

[2] British Medical Research Council Report, *Lancet*, 1948, i, 853.

; the concentration in the white corpuscles followed the same course, he final minimal level was rather higher than in (ii), *i.e.* about 2 mg./100 g.

ELATION BETWEEN BLOOD VITAMIN-*C* LEVEL AND CLINICAL STATE. symptoms first appeared after about 5 months in the vitamin-free group, ıtervals between the virtual disappearance of vitamin-*C* from the plasma white cells and the appearance of symptoms were respectively 14 weeks 3–6 weeks. It should be noted that although vitamin-*C* disappeared the plasma in the 10 mg. intake group, the subjects remained normal. virtual disappearance of -*C* from the *plasma* thus does *not* necessarily ate that scurvy is present or imminent; on the other hand a plasma above 0·1 mg./100 c.c. definitely excludes the diagnosis of scurvy. The cell -*C* content was significantly different in the -*C*-free group and that ving 10 mg. of -*C*; a white cell -*C* content below 2 mg./100 g. (if con- d) supports a diagnosis of scurvy.



ı. 653.—Changes in Vitamin-*C* (Ascorbic Acid) concentration in Plasma and White Corpuscles in Complete Experimental Vitamin-*C* Deficiency in Man. (After Crandon *et al.*, *New Eng. J. Med.*, 1940, *223*, 354.)

LINICAL CHANGES.—(1) *Negative Findings.*—Many negative findings be emphasized. There was no decrease in the red cell or white cell t, platelet count, or hæmoglobin concentration; vitamin-*C* lack therefore ıo adverse effect on red marrow activity (p. 194). There were no general lary changes; the resistance of the skin capillaries and bleeding time 58) were normal; there was no nose-bleeding; blood or red cells did ıppear in the urine; there was no occult blood in the fæces; the micro- c appearances of the capillaries in the nail bed and in the conjunctiva normal. Body weight was maintained and there was no complaint of ral pains or weakness.[1] Resistance to infection was not decreased.

In Crandon's case the diet was so unpalatable that the food intake was low and there loss of 25 lb. in weight in 6 months; his physical efficiency deteriorated and he do as much as the average octogenarian; a five minutes' walk was his peak effort. ıbly these changes were due to the prolonged period of general undernutrition . 1045).

(2) *Skin Changes.*—After 4–6 months there was enlargement and keratosis of the hair follicles in the skin. These changes were marked on the upper arms, back, buttocks, back of thighs, calves and shins. Microscopic examination showed that the follicles were plugged with horny material in which the hair was coiled or looped; subsequently the local vessels became congested and leaked; the follicles then became red and hæmorrhagic.

(3) *Gums.*—After 6 months, gum changes developed regularly, consisting of swelling and tiny hæmorrhages in the loops of the interdental papillæ. In two subjects who had some initial gingivitis gross changes occurred; the gums became very swollen and purplish in colour; later patches of necrosis appeared.

(4) *Wound Healing.*—Normal healing of a wound (*e.g.* that produced by a deep skin incision) involves development of new capillaries, proliferation of fibroblasts and the laying down of reticulin fibres followed later by the appearance of collagen fibres. In severe guinea-pig scurvy all these phases of wound healing are deficient. In Crandon, after 3 months' deprivation of vitamin-*C*, an experimental deep incision in the back healed rapidly and completely within 10 days. A similar incision performed after 6 months' deprivation showed good union of the epidermis after 10 days, but there were no signs of healing in the deeper parts of the wound which was filled with unorganized blood clot; no capillary or fibroblastic proliferation had taken place and there was no laying down of reticulin or collagen fibres (Fig. 654). There were considerable variations in the extent of impairment of wound healing in different subjects. The tensile strength of healed wounds was lower than normal. Scars of old wounds that had healed well in the early stages of vitamin-*C* deprivation became red or livid when clinical scurvy appeared.

(5) Neuromuscular coordination was slightly impaired and fatigue set in slightly more readily.

(6) In three of the 10 British volunteers who were completely deprived of vitamin-*C* grave complications developed which required immediate intensive vitamin-*C* medication; two developed signs of cardiac ischæmia, *i.e.* pain in the chest, dyspnœa, and serious electrocardiographic changes; one who may have had some initial affection of the spine developed changes in the vertebræ and a tuberculous paravertebral abscess.

There thus appear to be stages in the development of experimental scurvy: (i) an initial stage with changes in the vitamin-*C* content of the blood and no clinical signs; (ii) a stage of clinical scurvy consisting of changes in the skin, gums, and impaired wound healing; (iii) a final stage when life may be in immediate danger.

When 10 mg. of -*C* were administered daily to experimental subjects with well-developed scurvy, the skin changes cleared up in 1–2 months and the gums became normal in 3 months. When massive doses of -*C* were given recovery was rapid. Thus the chest pain and dyspnœa in the two subjects mentioned above were relieved in 24 hours. Crandon's experiment was ended by giving him 1000 mg. of vitamin-*C* intravenously daily; a wound inflicted on the final day of his treatment became completely healed in 10 days; the skin petechiæ faded in a week and the skin was quite normal in three weeks.

Renal Threshold for Vitamin-*C*.—When the plasma vitamin-*C* level

exceeds 0·85 mg./100 c.c., the vitamin appears in the urine. The presence of the vitamin in the urine is evidence that the subject is amply provided ("saturated") with the vitamin.

VITAMIN-C REQUIREMENTS IN MAN.—The experiments recorded above show that in adults, 10 mg. of vitamin-C daily are both completely protective and curative, even over periods as long as 14 months. To allow a margin of safety and to provide for individual variations a daily intake of 30 mg. of vitamin-C is probably desirable. There seems to be no clinically recognizable condition of sub-scurvy or pre-scurvy; people without the specific signs and blood changes of scurvy will not benefit in any way by supplementing the vitamin-C content of their diet.

CLINICAL SCURVY.—Scurvy occurs in *infants* who have been fed exclusively

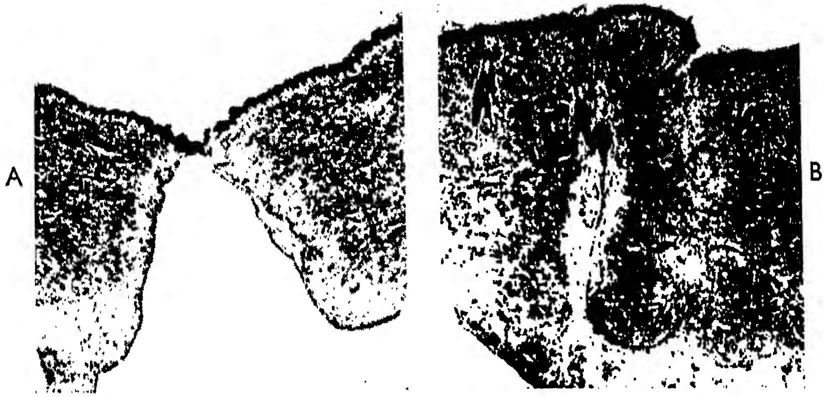


FIG. 654.—Delayed Wound Healing in Experimental Vitamin-C Deficiency in Man. (Crandon, Lund and Dill, *New Engl. J. Med.*, 1940, 223.)

A. and B. are biopsy specimens of skin.

A Wound made after 6 months' vitamin-C deprivation; note absence of healing. The skin is united superficially, but deeper tissue shows no healing whatever; the large space was occupied by an unorganized blood clot.

B Normal wound healing in same subject after 10 days' intravenous vitamin-C therapy.

on sterilized food, the vitamin-C content of which has thus been destroyed; the characteristic symptoms are *hæmorrhages* from various parts, and anæmia. After being fed for about 6 months on the defective diet, fretfulness and loss of appetite develop. Hæmorrhages occur (i) under the periosteum, *e.g.* at the lower end of the femur, giving rise to great pain, tenderness, and refusal to use the limb; (ii) from the gums if the teeth have erupted; (iii) from the kidney, intestine, and under the skin. There may be changes in the teeth themselves. The long latent period is noteworthy.

In *adults* similar symptoms of scurvy develop in people like sailors or prisoners who have been deprived for a long time of fresh food.

The clinical manifestations are thus far more severe than those observed in the controlled experiments. It must be remembered, however, that clinically the vitamin deficiency particularly in adults may be very *prolonged*, that multiple vitamin deficiencies may be present which may precipitate or aggravate the symptoms, and infections may develop which increase the rate of disappearance of vitamin-C from the body.

Mode of Action of Vitamin-*C*.—In experimental vitamin-*C* deficiency there is failure of certain tissues to lay down and maintain *normal intercellular ground substance* (including reticular and collagen fibres). This partly accounts for the poor wound healing and the reduced tensile strength of wounds. The *hæmorrhages* are attributed to *lack of cement substance in the capillary lining*. In animals with advanced experimental scurvy complex changes occur in the *bones*, especially at their ends; the bone cells are not buried in a firm matrix but are found lying in a liquid material. The grave terminal changes in experimental scurvy require some other kind of explanation.

PRINCIPLES OF DIETETICS[1]

Essentials of a Diet.—An adequate diet must have a *calorie* value sufficient to provide for the requirements of the basal metabolism, the stimulating action of the foodstuffs, loss in the fæces of energy-containing derivatives of the foodstuffs, and the needs of varying degrees of muscular work. It must have adequate amounts of *protein* (*essential amino-acids* (p. 881)), *fat*, *carbohydrate*, *water*, and *salts* (*ions*) in suitable proportions, and an ample *vitamin* content. In children adequate provision must be made for *additional tissue* formation; the special requirements of *menstruation*, *pregnancy*, *lactation* and *illness* must be taken into account. As the food has to be eaten by discriminating human beings it must be well cooked, look attractive, taste nice, and be reasonably varied from day to day.[2]

Energy Requirements.—A distinction must be drawn between the energy value of the food as purchased and of the food absorbed after digestion. A variable loss occurs during the preparation and cooking of the food; a further small percentage escapes digestion and absorption and is eliminated in the fæces, the loss being greater on a vegetable diet (because of its cellulose content (p. 815)) than on an animal diet. It is usual to deduct 10% from the theoretical calorie value of the mixed diet to allow for these losses. Thus, a ration with a theoretical yield of 1000 Cal. would only provide 900 Cal. in the body.

(i) The basal energy requirements can be calculated from the surface area; in an adult man it is 40 Cal. per sq. metre per hour; the average surface area is 1·8 metres; so the hourly basal energy output is 72 Cal. (p. 378). During 8 *hours of sleep* at basal level the total energy output is 576 Cal.

(ii) During waking hours the taking of food stimulates metabolism by 5–10%. During *8 hours of leisure* the ordinary small activities of daily life increase metabolism by about 40 Cal. per hour. An hour's walk at 2½, 3¾, and 5¼ miles per hour respectively involves an extra output of 140, 240, and 580 Cal.; an hour's dancing is equivalent to 240 extra Cal.; cycling, 175 Cal.;

[1] Jolliffe, Cannon, and Tisdall, *Clinical Nutrition*, N.Y., 1950. Davidson and Anderson, *Text Book of Dietetics*, 2nd edn., Edinburgh, 1947.

[2] "Nutrition is the science of test-tubes, gastronomy is the art of taste buds" (André Simon) (and, of course, of the olfactory nerve endings).

)xing, 800 Cal. The total calorie output in these 8 hours of leisure might .us be :

8 hours' basal (8 × 72)	= 576
Stimulating action of food	= 50
7 hours' small activities (7 × 40)	= 280
1 hour's exercise	= 240
Total	1146 Cal.

(iii) The energy output during *8 hours of work* varies enormously with the nd of work performed. The calories expended *per hour* above basal for some cupations are as follows: typing, writing, 40; sewing, 40; dish washing,); sweeping and cleaning, 100; laundry, 170; tailor, 80; carpenter, 150; etal work, 180; blacksmith, 300; stonemason, 330; sawing wood, 400; al mining (cutting), 120, timbering, 200. The 8 hours' additional energy itput can be classified thus: sedentary work, under 400 Cal.; light work,)0–700 Cal.; moderate work, 700–1100 Cal.; heavy work, 1100 Cal. or over ven the most concentrated brain work causes no measurable increase in iergy output.

The energy output for the 8-hour work period might be thus:

8 hours' basal (8 × 72)	= 576
Stimulating action of food	= 50
Energy value of work	= 400 up to 2000 Cal. or more.

(iv) The *total* daily output for a *sedentary* worker (work expenditure=50 al. per hour) might thus be:

8 hours of sleep . . .	576
7 hours of leisure + 1 hour's exercise	1146
8 hours of work (626 + (8 × 50)) .	1026
Total . .	2748 Cal.

CHILDREN.—The energy needs of children are difficult to compute. The asal metabolism is considerably greater per sq. metre of surface in children han in adults (p. 378). The extra energy demanded by the *growth* processes ; very small. Between the ages of 11 and 16, both sexes put on weight at he average rate of about 4 kg. a year, which is equivalent to the negligible [uantity of 30 Cal. per day.

It is still more difficult to compute the very variable energy output due o muscular *activity* in children. Probably any reduction of the food intake auses a diminution in activity, before it affects the processes of the body vhich are essential to health and growth.

Standards of Calorie Requirements.—In view of the enormous)ractical importance of the subject, not only from the physiological but also 'rom the social and political standpoint, the energy requirements as worked)ut by an expert Committee of the British Ministry of Health are quoted)elow. They laid down that the *average caloric requirement of a man doing ight work is* 3000 *Cal. net* (*i.e. available after absorption*), *or* 3300 *Cal. as)ought.* The average housewife needs about 10% less, *i.e.* 2700 Cal. net; a

woman doing more active work needs as much as the average male. These allowances probably err on the generous side. During the last war the average food intake in Britain was less than these recommendations. Men doing hard work may need up to 4000 Cal. and more in appropriate cases. The needs of children are laid down as follows :

Age . . .	1–2	2–3	3–6	6–8	8–10	10–12	12–14
Calories . .	1000	1250	1550	1850	2150	2550	2900

Note that children over 12 require as much food as adults. For girls between 14 and 18, 2800–3000 Cal. are advised; for boys of the same age, 3000–3400 Cal., *i.e.* rather more than the requirements of an adult light worker.

Results of Inadequate Caloric Intake.—1. Moderate Calorie Deficiency. The Carnegie Experiment.—Twelve healthy subjects whose normal food intake was equivalent to 3100 Cal. per day were put on a diet containing only 1600–1800 Cal. for 5 weeks; the diet was adequate in other respects. During this period they lost 10% of their body weight. The food intake was then increased to 1970 Cal. daily; on this diet the body weight remained approximately stationary over a period of several months. The striking economy in food intake was achieved, however, at the cost of a considerable deterioration in the general physical and mental state. The subjects felt weak and tired easily; they were unable to work hard or for a long time; sexual drive was decreased. Their listless depressed state led to a saving on the 500 Cal. usually needed for "small activities" and exercise (p. 1043). The chronic state of undernutrition decreased the basal metabolic rate as explained on p. 378. As far as they could they reduced their energy output further by doing less work. The energy intake provided no reserve for the additional requirements of fever or illness. The final picture was similar to that found among the poor in many backward countries. As their diet is inadequate these people are feeble and unable and unwilling to work hard; they have no drive or initiative. Owing to their depressed physical state their productivity falls off, and in agricultural communities the food output declines. A vicious circle is thus established in which malnutrition leads to weakness which in its turn aggravates the degree of malnutrition. Any illness which decreases the energy intake or increases the energy loss may have serious consequences. Such a population is unlikely to be capable of sustained vigorous physical or moral resistance against an oppressor. If external aid were available to produce an initial improvement in the nutritional state the resulting increase in strength might so improve productivity that a satisfactory diet might ultimately be made available by the people's own efforts.

2. Severe Calorie Deficiency. The Minnesota Experiment.—The effects of more severe and prolonged semi-starvation were studied by a group in the University of Minnesota.[1] Thirty-two healthy young adults whose

[1] Keys *et al.*, *Biology of Human Starvation*, Minneapolis, 1950. This monograph gives an encyclopædic survey of the whole subject of undernutrition.

average control intake was 3490 Cal. were kept for 24 weeks on a diet of 1570 Cal. containing approximately 50–55 g. of protein, 30–45 g. of fat and 300 g. of carbohydrate; as a result they lost 25% of their original body weight. The intake of vitamins and of minerals was adequate.

General Changes.—The wasting affected the face and the body alike; the emaciated features, the protruding ribs, the unprotected ischial tuberosities and the badly fitting clothes and shoes made the subjects acutely aware of the degree of physical deterioration that had taken place. The skin became rough, thin and pigmented; it was generally cold and slightly cyanosed. The hair grew more slowly and fell out in large amounts; the nails also grew more slowly. Œdema was common after the twelfth week and was specially noticeable in the ankles, knees and face. Cuts and wounds bled less freely and healed more slowly. Tolerance to heat was increased but that to cold was diminished. The subjects liked their food and drink to be served unusually hot. They constantly complained of their cold hands and feet; even in the hot summer they wore extra clothes by day, and covered themselves with many blankets at night. The blood pressure was unchanged but the resting pulse rate was very low (35 per minute). They complained of giddiness or momentary blackout on getting up from the sitting or lying position; actual fainting was rare. There was no impairment of visual acuity but many subjects complained of transient visual disturbances such as inability to focus, eye-aches, and "spots" before the eyes. There was a slight increase in auditory acuity; it is uncertain whether the frequent complaint that ordinary sounds and noises were disturbing and annoying was the result of the "improved" auditory sensitivity or was due to a state of "irritability." Muscle cramps and muscle soreness were common; the extremities often went "to sleep." Voluntary movements were noticeably slower; the subjects felt weak and tired easily; they moved cautiously, climbing stairs one step at a time. Coordination was disturbed; the men sometimes tripped over the kerb and bumped into objects which they intended to avoid. There was a reduction in self-initiated, spontaneous activities; as starvation progressed, fewer and fewer things stimulated the men to overt action. They said that their increasing weakness, loss of ambition, narrowing of interests, depression, irritability and loss of libido made them feel they were growing prematurely old.

Quantitative Changes.—Some of the principal quantitative changes observed are summarized below; in all cases the results quoted are the mean values for the group as a whole.

	Control.	*After 24 weeks' semi-starvation.*
Calorie intake	3490	1570
Body weight (kg.)	70	52·5
Body fat (kg.) (calculated)	9·8	3·0
Hæmatocrit value (%)	46·4	36·6
Hæmoglobin (g–%)	15·1	11·7
Total blood volume (c.c.)	5850	5350
Total plasma volume (c.c.)	3130	3390
Blood volume (c.c./kg.)	84	101

	Control.	After 24 weeks' semi-starvation.
Plasma volume (c.c./kg.) .	45	64
Total plasma protein (g./100 c.c.)	6·7	6·0
Plasma albumin (g./100 c.c.) .	4·3	3·9
Plasma cholesterol (mg./100 c.c.)	169	151
Basal pulmonary ventilation (litres/min.) . . .	4·8	3·3
Vital capacity (litres) . .	5·2	4·9
Basal O_2 consumption (c.c./min.)	228	139
Basal O_2 consumption (c.c./sq.M./min.) . .	122	84
Basal pulse rate (beats/min.) .	55	35
Systolic blood pressure (mm. Hg.)	106	99
Pulse pressure (mm. Hg.) .	36	31
ECG :		
height of $T_1+T_2+T_3$ (mV×10)	5·6	3·3
height of R in Lead I (mV×10)	3·7	1·4
height of R in Lead II (mV×10)	9·9	6·2
Time for which work test could be kept up (seconds) .	242	50
Strength of hand grip (kg.) .	58	42
Back dynamometer test (kg.) .	150	105

Some of the results noted above deserve further consideration.

(1) *Body Weight.*—The loss of body weight was masked to some extent by the presence of œdema. On refeeding, an initial further loss of weight might occur owing to loss of fluid. The muscles wasted rapidly and the changes in the flaccid muscles contributes markedly to the apathetic mask-like expression of famine victims. In some cases the œdema gave the face a bloated expression with swollen eyelids and puffy cheeks.

(2) *Endocrine Glands.*—There was marked atrophy of the thyroid, evidence of dysfunction of the pituitary and changes in the adrenals.

(3) *Body Fluids. Œdema.*—The œdema was associated with a fall in the plasma protein concentration and osmotic pressure, but it was not regularly related to the extent of these plasma changes (cf. p. 115). The total plasma volume was little altered, but was greatly increased relative to the reduced total body bulk; there was probably a relative excess of body fluid; the mechanism of this fluid retention is obscure, but then little is known about the factors which normally regulate the total volume of body fluid. In spite of the œdema, polyuria was often a marked symptom associated with a correspondingly high fluid intake. The fall in hæmatocrit value and hæmoglobin level was due to decreased hæmatopoiesis (cf. pp. 163, 210).

(4) *Circulation.*—The remarkable degree of bradycardia should be emphasized; it may be due in part to the presence of thyroid deficiency. The peripheral blood flow was reduced and the venous pressure lowered; the blood pressure was unchanged.

(5) Among the *alimentary* symptoms were flatulence, colic and diarrhoea. Gastric tone and motility were depressed.

(6) *Metabolic Rate.*—The B.M.R. was reduced whatever criterion was employed; there was a decrease in the B.M.R. per kg. of 20%, per sq.m. of 30%, and per man of 40%. The fall in resting metabolism was due only in part to the smaller mass of active tissue; to some extent it may be due to thyroid deficiency. The average saving of energy produced in this way was of the order of 600 Cal. daily.

(7) *Negative Nitrogen Balance.*—There was a consistent state of negative nitrogen balance for the reasons explained on p. 901. The relationship of calorie intake to N balance is well shown in the Table below (after Boothby and Bernhardt).

		Food Intake.		*Nitrogen Balance.*	*Change in Body Wt.*
Period.	*Date.*	*Calories.*	*N (g.)*	*g.*	*kg.*
I	2nd June–15th June	894	9·4	−3·4	−2·9
II	16th June–1st July	1296	10·1	−2·1	−0·3
III	2nd July–19th July	1620	11·6	−1·2	−0·2
IV	24th July–12th Aug.	3353	15·2	+2·0	+2·3

(8) *Sensation of Hunger.*—In people totally deprived of food the feeling of hunger disappears after a few days; in semi-starved people on the other hand the hunger sensation becomes progressively accentuated. It is known that strong contractions of the empty stomach give rise to a sense of hunger (hunger pains), but it is not likely that hunger is always or generally produced in this way. Thus these subjects often complained of hunger immediately after a large meal and when gastric motility and tone were depressed. The cause of the hunger sense is as obscure as that of thirst (cf. p. 68).

REHABILITATION.—After the fast, the subjects were placed on progressively larger diets until full recovery occurred. Recovery from dizziness, apathy and lethargy was rapid; tiredness, loss of sex drive, and weakness improved slowly. The œdema gradually disappeared though in some men there was initially little change or even an increase in œdema. Cramps and vague aches and pains persisted for some time. Some of the men had new complaints such as flatulence, distension, and stomach-ache. Those subjects who gained most weight became concerned about their increasing sluggishness, general flabbiness, and the tendency for fat to accumulate on the abdomen and buttocks. At the end of three months even the subjects on the highest calorie intake were still in a poorer physical condition than in the control period. It was not until after an additional 3 months of "normal living and supernormal eating" that the physical capacity approached pre-experimental levels.

It has been generally thought that patients subjected to prolonged severe starvation cannot swallow, digest or absorb simple foods. During the Dutch post-war famine elaborate preparations were made to feed the patients by stomach tube or intravenously with protein hydrolysates and glucose solutions. Experience showed that those who were too ill to take food by mouth died whatever was done to them; but the others had a good appetite, could take normal food almost from the first and were able to digest and absorb it well if it was given in small quantities at frequent intervals. They tolerated separated milk and even large amounts of butter-fat quite well.

In clinical rehabilitation it is generally advisable to concentrate on increasing the calorie intake by giving any foods immediately available,

which in Europe were grains and potatoes; these foods also supply additional protein and vitamins.

3. **Clinical Undernutrition.**—In every age and in every land people have starved and the record of the 20th century in this respect is as bad as most. Hunger is generally due to ignorance, misfortune or human malevolence; of the last we have experienced more than our fair share. The effects of undernutrition vary with its nature. In the recent European famines the main deficiency was in calories, as the wholemeal bread and vegetables eaten (especially potatoes) prevented serious protein or vitamin deficiencies. Vitamin requirements (especially of the -*B* group) are lower when the calorie intake is reduced. In the Eastern famines, of which the Singapore captivity is an excellent example, the effects of vitamin deficiency often predominated (p. 1031).[1] In clinical studies the immense importance of the psychological changes must be emphasized; the starved person differs in mind as well as in body from the well-nourished person.

Starved people differ in the degree and kind of the undernutrition and in the effects of the associated brutalities or other complications. The results have been reviewed at length in the monograph by Keys and his associates. Some recent clinical studies are summarized below.

(1) Observations [2] on Prisoners of War in World War II.—The changes in the badly fed soldier in the prison camp were as follows: "First there was the loss of the natural feeling of well-being. A growing feeling of hunger followed and gradually increased in intensity until, after about 3 weeks, the whole thought of the prisoner was concentrated on his food." This insistent feeling of hunger increased with time; the half-starved man would go to great lengths of ingenuity and dishonesty to obtain small amounts of extra nourishment. Only when death was imminent did the desire for food slowly vanish.

"The next notable abnormality was a progressive mental and physical lethargy. The desire for sleep increased; the number of hours that an adult male would wish to remain in bed, partly dozing but for the most part in genuine deep sleep, steadily rose from the normal 8 hours to 16 or more out of the 24. Finally, the only way to rouse a man from his bed was by the mention of food." Shortly after this came rapid fatigue during mental or physical effort. The experienced card-player, for instance, would forget the cards that had just been played.

In the fully developed picture of semi-starvation the slowness of movement was notable. The gait in walking was a shuffle and suggested a slow-motion picture. Speech was slow, and obtaining responses to questions was tedious. Their slow movements and delayed reaction times made these men especially prone to accidents: "Though they could see that a fall of rock or branch of a tree was imminent, their actions were too slow to move out of danger." The desire for tobacco greatly increased and seemed stronger than the desire for food, since men would barter even the little food they had for a small amount of tobacco. Standards of cleanliness deteriorated drastically, and pride in personal appearance was lost. Normal moral standards lost all influence on the behaviour of the starving men and they would steal food from

[1] The results of lack of individual dietary constituents, *e.g.* iron, iodine, vitamins, are reviewed in the appropriate sections in this book.

[2] Leyton, *Lancet*, 1946, ii, 73.

their best friends or barter their clothes for tobacco. "None of the other hardships suffered by fighting men observed by me [Leyton] brought about such a rapid or complete degeneration of character as chronic starvation."

(2) THE GERMAN EXTERMINATION CAMPS.[1]—"The psychological atmosphere of the extermination camps is beyond our capacity for empathy. The official daily ration in the notorious Auschwitz Camp was one litre of watery soup, 250 g. of bread and about 25 g. of margarine or sausage or imitation honey, providing an estimated maximum of 1000 Cal. per day. The German camp doctors did not expect the inmates to survive much beyond 6 months. The camps were grossly overcrowded, with 500 to 1000 people herded together in blocks of 100 feet by 20 feet. Sanitary facilities were utterly inadequate. There was a shortage of water. The inmates were infested with lice, and many had scabies. They were under constant threat of being subjected to the brutalities of the guards. But dominating all the thoughts and fears was the dread death by gassing or burning. The extermination verdict was entered in the camp books under the euphemistic term of 'transfer for special treatment'."

(3) THE BELSEN CAMP, 1945.—Belsen was a detention camp, not an extermination camp, *i.e.* the prisoners there were allowed to die but were not usually killed. The diet was always inadequate, but in the winter of 1944–45 it became grossly deficient. Lipscomb [2] described what he found when the camp was taken over by the British Army in April 1945. At this time the camp was estimated to contain about 50,000 people, with some 10,000 lying dead in the huts and around the camp. "The most conspicuous psychological abnormality was a degradation of moral standards characterized by increasing selfishness, and it was more or less proportional to the degree of undernutrition. In the first stage consideration for others was limited to personal friends, then the circle contracted to child or parent, and finally only the instinct to survive remained. Emotional response became progressively lowered and consciousness of sex was lost. Eventually all self-respect disappeared and the only interest left was to obtain something which could be eaten, even human flesh. Even among those not grossly undernourished there was a blunted sensitivity to scenes of cruelty and death."

Cases with mental disorders were surprisingly few in view of the harrowing experiences, the physical suffering, and the fear of hunger, torture, and death to which the internees were exposed.[3]

[1] The quotation is from the monograph by Keys *et al.*, *Biology of Human Starvation*, 1950.

[2] Lipscomb, *Lancet*, 1945, ii, 313.

[3] But semi-starvation may occur even in England. The following is an extract from Howard's description (quoted from Keys' monograph) of the results of hunger in London during the winter of 1837–38 when unemployment was general. "The first indications of a deficiency of food are languor, exhaustion, and general debility, with a distressing feeling of faintness and sinking at the præcordia, chilliness, vertigo, and a tendency to syncope, unsteadiness of movements, the voice weak and tremulous. Mental powers exhibit a languor and dullness proportionate to the degree of physical debility. The sufferer is listless and depressed and often manifests a remarkable apathy to his condition." In more advanced cases "the feeling of prostration becomes quite overpowering, and the exhaustion and muscular weakness so great, that the erect posture can, with difficulty, be maintained. The depression of all the vital and mental powers is fearfully augmented. Dizziness, transient dimness of vision, staggering, and syncope are common and very apt to be produced by the erect posture. Notwithstanding general languor, however, the patient sometimes manifests a highly nervous state; he is startled by any sudden voice and worried by most trifling occurrences." The findings are almost identical with those of the Minnesota experiment and are evidence of the severity of the starvation.

4. Complete Starvation.—See p. 900.

Composition of Foodstuffs.—Before dealing with the proportion of protein, fat, and carbohydrate that should be employed, the composition of the common foodstuffs must be surveyed (see Tables, pp. 1052–1054).

Constitution of Normal Diet.—The constitution of an average normal British diet is usually given as follows: Protein, 100 g. (410 Cal.); fat, 100 g. (930 Cal.); carbohydrate, 400 g. (1640 Cal.). Total calories=about 3000.

1. **Protein.**—Protein is an indispensable constituent of the diet because it is the only source of the *amino-acids* including the *essential* amino-acids which cannot be synthesized in the body, (p. 881). Amino-acids are needed to build up new tissue during the period of growth or pregnancy and to provide the milk proteins during lactation. They are necessary to maintain the structure of every tissue cell including its content of protein-containing enzyme systems; to provide the raw materials for the manufacture of certain external secretions (digestive enzymes of alimentary canal) and internal secretions (*e.g.* those of the anterior pituitary, thyroid, adrenal medulla); and to maintain the normal concentrations of plasma proteins and hæmoglobin. As explained on p. 879, there is normally a *dynamic equilibrium* between the amino-acids derived from the food and those derived from the breakdown of the tissues, and also between the amino-acids derived from one tissue and those derived from another. But in the course of this endless flux some amino-acids are *broken down irreversibly* and their nitrogen content is excreted in the urine mainly as urea. The minimum protein intake in the adult must contain enough of *each essential amino-acid* to make good this steady loss. The *individual* proteins of *animal* tissues generally closely resemble those of human tissues in their amino-acid composition; they can consequently be employed more economically for repair and growth; those with such a composition are called proteins of *high biological value*. The *individual* proteins of *vegetable* foods frequently have a very different type of amino-acid pattern and so cannot be so economically built up into human tissues; such proteins are said to be of low biological value. But it must be pointed out that this discussion is quite academic; we do not eat individual animal or individual vegetable proteins; we eat *foodstuffs*, *e.g.* meat, potato, bread, egg, milk, all of which consist of a *mixture of several proteins*. The important thing to find out is what is the amino-acid composition and the biological value of the *protein mixture in a single foodstuff* or more important still in the *mixture of foodstuffs found in typical human diets* (cf. Fig. 563, p. 881; Fig. 578, p. 903).

There are two ways (apart from chemical analysis) of testing the biological value of an individual protein: (i) to see whether it can maintain satisfactory growth and health when it is the sole source of protein supplied; (ii) to determine the minimum amount of the protein which must be added to a diet adequate in other respects to maintain nitrogenous equilibrium. Similar determinations can be carried out on a mixture of proteins, a foodstuff or a mixture of foodstuffs. Using these criteria, certain *isolated* plant proteins, such as gliadin or zein, are very inadequate; the proteins of flour made from wheat endosperm are not satisfactory (p. 1058); on the other hand, the *mixed proteins* of wheat or maize as found in high extraction flour, by compensating for one another's deficiences or quantitative peculiarities, maintain nitrogenous equilibrium at a fairly low level, particularly if

COMPOSITION AND CALORIE VALUE OF PRINCIPAL FOODSTUFFS.[1, 2]

FOOD	Available				Minerals in mg. per 100 g. (Divide values by 4 to get approx. content per oz.)			Vitamins: values in international units (i.u.) or mg. per 100 g. (Divide values by 4 to get approx. content per oz.)			
	Protein per cent.	Fat per cent.	Carbo-hydrate per cent.	Calories per 1 oz. (28 g.).	Fe	P	Ca	-A (i.u.)	-B_1 (i.u.)	-C (mg.)	-D (i.u.)
FRUIT:											
Apple, eating	0·3	..	12	14	0·2	7	4	75	25	10	..
Banana	1·1	..	19	23	0·4	28	7	300	25	10	..
Grapes	0·6	..	16	19	0·3	21	19	50	+	3	..
Orange	0·8	..	9	11	0·3	23	41	75	25	50	..
VEGETABLES:											
Beans, baked	6·0	0·4	17	28	2·1	184	61	..	100	..	..
Beans, haricot, boiled	6·6	..	17	27	2·5	122	64	..	..	..	..
Cabbage, winter, boiled	0·8	..	1	2	0·5	16	58	50	60	50	..
Lettuce	1·1	..	2	3	0·7	30	25	1,500	15	15	..
Onions	0·9	..	5	7	0·3	30	31	+	20	5	..
Parsnips, boiled	1·3	..	14	17	0·5	30	65	+	50	15	..
Potatoes, old, boiled	1·4	..	20	24	0·5	30	4	+	++	+	..
Spinach, boiled	5·1	..	1	7	4	90	600	12,000	30	27	..
Swede, boiled	0·9	..	4	5	0·3	18	41	..	..	17	..
Tomatoes	1·0	..	4	4	0·5	25	15	1,000	80	20	..
Turnip tops, boiled	3	..	..	3	3	45	100	6,000	25	20	..
Watercress	3	..	1	4	1·6	52	220	1,500	45	60	..
FISH:											
Bloater, grilled	16·7	13	..	53	1·6	263	91	..	..	..	..
Cod, boiled	14·6	0·7	..	19	0·5	196	12	..	..	..	..
Haddock, boiled	16·7	0·6	..	20	0·5	178	41	..	..	..	..
Herring, soused	15·5	12	..	48	1·5	300	53	..	..	..	..
Kipper, baked	13	6	..	31	0·8	230	35	..	..	..	..
Plaice	9·8	1	..	14	0·3	133	20	..	..	..	..
Salmon, canned	19·1	6	..	39	1·3	285	66	30–600	25	..	500
MEAT:											
Bacon, fried	27	35	..	120	4	236	23	..	80	..	..
Beef, lean and fat	21	32	..	108	4·6	236	6	50	50	..	..
Ham, lean	23	13	..	60	2·6	244	17	..	150	..	..
Ham, lean and fat	16	40	..	122	2·5	190	13	..	200	..	..
Mutton, leg, roast	25	20	..	82	4	240	4	..	++	..	..
Pork, loin, roast	19·5	40	..	123	2·3	185	7	..	200	..	..
Veal, roast	30·5	11	..	64	2·5	355	14	..	20	..	..
Beef, corned	22	15	..	64	9·8	120	13	..	6	..	..
CEREALS:											
Bread, white	7·9	0·7	54	74	1·0	73	23	25	12	..	..
Bread, brown	8·4	1·6	44	65	2·7	213	30	25	50	..	..
Biscuits, digestive	10·5	20	66	143	1·6	134	44	..	..	..	..
Flour, white	10·7	1	80	108	0·9	102	18	..	10	..	..
Oatmeal, dry wt.	13·3	9	73	123	4	380	55	15	90	..	..

1 i.u. of $-B_1 = 3\ \mu g$.

1 Table prepared by W. F. Floyd and I. Calma.
2 M'Cance and Widdowson, *Chemical Composition of Foodstuffs*, Med. Res. Council Sp. Rep. No. 235, 1940; Hawley and Maurer-Mast, *Fundamentals of Nutrition*, 1940.

COMPOSITION AND CALORIE VALUE OF PRINCIPAL FOODSTUFFS (*continued*)

FOOD	Available				Minerals in mg. per 100 g. (Divide values by 4 to get approx. content per oz.) [1]			Vitamins: values in international units (i.u.) or mg. per 100 g. (Divide values by 4 to get approx. content per oz.)			
	Protein per cent.	Fat per cent.	Carbo-hydrate per cent.	Calories per 1 oz. (28 g.)	Fe	P	Ca	-*A* (i.u.)	-B_1 (i.u.)	-*C* (mg.)	-*D* (i.u.)
MISCELLANEOUS:											
Milk	3·3	4	4·6	20	0·08	95	120	150–600	30	2–50	2–20
Cream ("20 per cent.")	2·4	18	1	50	0·2	25	59	1,000–15,000	20	..	35–50
Butter	0·4	85	..	226	0·2	24	15	1,000–6,000	40	..	25–50
Margarine [2]	0·2	85	..	226	0·3	12	4	..	..	..	..
Cheese, cottage	25	35	..	120	0·6	545	810	100	..	..	..
Cheese, cream	3·2	86	..	232	0·1	44	30	3,000–5,000	..	..	..
Chocolate, milk	7·4	34	52	160	1·7	215	175	..	..	..	..
Chocolate, plain	4·6	32	59	160	3·3	140	26	..	..	..	..
Cocoa	20·4	26	35	126	14·3	685	51	..	..	..	..
Egg (=2 oz.)	12	12	..	46	2·5	218	56	1,000–4,500	30	..	150
Sugar	..	..	100	116	..	..	..	..	..	..	..
Golden Syrup	0·3	..	79	92	1·4	20	26	..	..	..	..
Honey	0·4	..	76	90	0·4	17	5	..	..	..	..
Jam	0·5	..	69	81	1·2	18	15	..	..	..	..
Marmalade	..	..	69	81	0·6	12	35		..	..	..

[1]

Mg./100 g.	Fe	P	Ca
Fair content	0·5–2	50–200	50–200
Good content	2.0–5·0	200–500	200–500
Very good content	> 5·0	> 500	> 500

[2] "Unfortified Margarine."

they are supplemented by small quantities of other proteins such as those of milk or meat. *In any mixed diet, even if wholly of plant origin, the proteins are sure to be sufficiently varied to compensate for any individual inadequacies in amino-acid content, if only the total amount of protein is sufficient.* The recommendation for adults is that the protein intake should not be less than 1 g. per kg. body weight, *some of it* in the form of animal protein. In the British Rationing System a minimum intake of 30 g. of animal protein daily was guaranteed. Pregnant and lactating women need more, *e.g.* 1·5 to 2 g. per kg. of total protein and (for safety) a larger amount of animal protein (p. 1059).

Nitrogenous equilibrium and unimpaired health and mental and physical vigour have been maintained for many months on diets which contained no more than 30–50 g. protein daily. But the proteins were specially selected and were of high biological value and the experiments were carried out on a very small number of subjects for too short a period. In some experiments nitrogenous equilibrium was maintained on 30–40 g. of protein derived exclusively from vegetable sources (cereals, potatoes, and other vegetables and fruit). To provide a margin of safety a *minimum of* 80 *g.* of protein should be provided in the form of a *varied diet.* Bayliss said, "Take care of the calories and the proteins will take care of themselves," because if dietaries are studied it is found that when the calorie value of a mixed diet is adequate it generally contains 100 g. of protein or even more.[1]

2. **Fat.**[2] The amount of fat consumed varies with the country, economic status, occupation, and the general circumstances. The maximum fat content of the really native diets in Japan before the second world war was thought to be about 30 g. which was then the European minimum. The amount of fat eaten increases steadily with the income (in an uncontrolled economy); the Inter-Allied Food Commission adopted 57 g. daily as the minimum fat ration during the first world war.

Experiments of short duration have been carried out in which health and weight were maintained on diets of high total calorie value containing as little as 10–14 g. of fat daily. Although the body can adapt itself to partial or even complete deprivation of fat for short periods, deleterious effects may become apparent after months or years.

The significance of fat in the diet depends on several factors. Firstly, it is almost completely absorbed from the alimentary canal. The bulk of the food becomes of importance when the total energy requirements of the body are very large. Weight for weight, fat has double the calorie value of starch or sugar. Fat, in addition, is taken without admixture in a pure form,

[1] Proteins of high and low biological value are sometimes called "first class" and "second class" proteins respectively. As the best known *individual* proteins of low biological value are derived from vegetable foods there has been a regrettable tendency to equate the terms "first" and "second" class generally with animal and vegetable proteins respectively. Such misuse of the terms should be strenuously avoided. Before vegetable proteins in general are dismissed as second class it should be remembered that many animals build up their muscles (*i.e.* meat) from the proteins of the humble grass. Nebuchadnezzar lived for a time on grass and so it is claimed have exceptional individuals since his time. Many people who have been accustomed to a diet rich in animal food feel unhappy without it and their efficiency as workers may fall off as a result; but it is hard to assess the relative importance of physiological and psychological factors in such cases.

[2] Anderson and Williams, *Physiol. Rev.*, 1937, *17*, 33.

whereas the other foods are all mixed with a considerable proportion of water; when starch is cooked it is swollen up with five to ten times its volume of water. As has been pointed out, the Swedish and American lumbermen and the Welsh miners obtain a large part of their huge caloric intake from fat. Carbohydrates are more subject to fermentative changes in the intestine, with the production of gases and general discomfort. It is advisable that the normal diet of 3000 Cal. should contain at least 75 g. of fat (700 Cal.); the fat content of the diet should always be raised when there is a large increase in the energy expenditure of the body, either in the form of work or because of exposure to cold.

The *animal* fats are the most important sources of some of the vitamins, *i.e.* -*A* and -*D* (pp. 1021, 1009). The vegetable fats are an equally effective source of energy, but are deficient in vitamins, except -*E* (p. 1086). When butter is rationed it is essential from the point of view of the health of the community that margarine should have vitamins -*A* and -*D* added to make it equivalent to butter in its vitamin content. This has been enforced in Great Britain. *Adequately vitaminized margarine* (made from vegetable oils) is equal dietetically in all respects to butter, and is generally much cheaper.

3. **Carbohydrates** furnish more than 50% of the energy content of most diets, and are a cheap and readily obtained food. If the amount of carbohydrate ingested is greatly reduced in amount ketosis may develop (p. 875). As both carbohydrate and fat serve chiefly as sources of energy, they can replace one another to a considerable extent, so long as precautions are taken to ensure the minimum amounts of fat specified above.[1]

4. The **mineral constituents** of the human body amount to 4·3–4·4%.[2]

(i) The only salt commonly consumed as such is sodium chloride; it is also present in small amounts in the food, particularly in milk and in vegetables. The minimum requirements are 1–2 g. of NaCl daily; while custom varies considerably, the average intake is about 8–10 g. The effects of NaCl deficiency are considered on p. 64.

(ii) 0·9–1 g. of *calcium* is needed daily; the minimum necessary for the maintenance of a calcium balance is 0·63 g. of calcium per 70 kg. of body weight. A sufficient calcium supply is very important, especially in children, and is best obtained by the ingestion of liberal quantities of milk and "fortified" bread (cf. pp. 997 *et seq.*)

(iii) 0·88 g. of *phosphate* is the minimum needed per 70 kg. body weight, but there is no risk of phosphorus shortage in a diet yielding 3000 Cal. daily unless large amounts of white flour are consumed as the principal cereal. (See pp. 1001 *et seq.* for main sources of phosphorus.)

(iv) The daily intake of *iron* should not be less than 12 mg., an amount which should be increased in pregnancy and in lactation (cf. pp. 204, 1059 *et seq.*). The rôle of *copper* is considered on p. 213.

(v) *Iodine* in minute traces is essential for the formation of thyroxine, the active principal of the thyroid gland. In districts in which the iodine supplies in the drinking water are insufficient, simple goitre frequently develops. This may be prevented, or the condition may be cured, by the use of iodized table salt, or by the deliberate addition of iodine to the water supply (p. 987).

[1] For relative value of white and brown sugar, see Cole, *Brit. med. J.*, 1939, ii, 340.
[2] Shohl, *Mineral Metabolism*, American Chemical Society, 1939.

5. **Vitamins.**—The detailed distribution of the vitamins is given in the special sections devoted to them (pp. 151, 1008, 1021, 1024, 1038, and 1086). For practical purposes the following generalizations serve as an adequate guide during normal peace-time conditions. Fresh fruit juices and green vegetables ensure an adequate supply of vitamin-*C*. High (85%) extraction flour goes a long way to providing all the necessary vitamin-*B* complex. A plentiful supply of animal fats (milk and milk products, meat fat, and vitaminized margarine) and eggs gives all the necessary vitamin-*A* and -*D* for adults; green vegetables and carrots are important sources of -*A*; in the case of children, *cod-liver oil* or *halibut-liver oil* should be added as a supplementary source of these vitamins. Vitamin-*E* or -*K* requirements need not be considered in preparing dietaries.

For rôle of *extrinsic factor* in blood formation see p. 200.

Bread.[1]—As bread is the most important single constituent of the diet in many countries, its composition deserves detailed consideration. Fig. 655

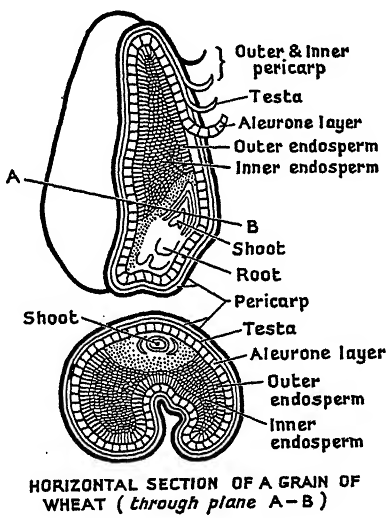


FIG. 655.—Structure of Wheat Grain. (McCance, *Lancet*, 1946, i, 77.)

shows a longitudinal and horizontal section through a wheat grain; the main divisions of the grain are as follows:

(i) *Bran*, consisting from without inwards of the outer and inner pericarp and the testa (these are fibrous and constitute 4% by weight of the grain) and the aleurone layer (8% of total weight).

(ii) *Endosperm*, divided into an outer layer (2%) and an inner part which constitutes 83% of the grain-weight.

(iii) *Germ*, divided into the embryo (1·2%) and the scutellum (1·5%).

The composition of whole English wheat, and of the endosperm and germ, and of 85% and 75% extraction flour, are shown in the Table on p. 1058.

[1] McCance, *Lancet*, 1946, i, 77.

	Whole Wheat.	Regions of Wheat grain.				Flour.	
		Endosperm.		Germ.		85% Extraction.	70% Extraction.
		Inner.	Outer.	Embryo.	Scutellum.		
Protein . g./100g.	8·9	8·1	14·6	33	29	8·55	7·92
Fat . . g./100g.	2·2	0·7	—	15	30	1·46	1·04
Carbohydrate g./100g.	66·8	75·8	—	—	—	72	74·5
Fibre . g./100g.	—	—	—	—	—	0·42	trace
P . . mg./100g.	311	59	1017	1160	1900	—	—
Phytate P. mg./100g.	213	8	874	400	1300	—	—
Fe . . mg./100g.	3	0·5	12	9	—	2·22	1·4
Ca . . mg./100g.	—	—	—	—	—	24·5	18·9
Vitamin-*B* units/g.	1·6	0·1	1·8	3	60	0·7	0·25
Riboflavin. μg./g.	1·5	0·4	1·8	15	—	1·2	0·6
Nicotinic Acid μg./g.	42	5	184	60	—	10	7·5

Notes.—1. The sign — means no data given.
2. Canadian (Manitoba) wheat contains 13·6 g–% of protein.
3. The aleurone has the same composition as the outer endosperm.
4. The germ is rich in vitamin-*E*.

Effects of Varying Extraction Rate.—(i) When wheat is milled varying proportions of the grain are retained. With 100% extraction the entire grain is ground to flour; with 85% extraction most of the bran is rejected but most of the germ is retained; with 70% extraction all the bran and germ is lost and the flour is derived from the inner endosperm alone.

(ii) Seventy per cent. extraction flour is white, 80% is off-white, 85% is a shade darker, 100% is brown.

(iii) Bran contains 20% and the germ 4·5% of the total grain proteins. The germ proteins are as valuable nutritionally as the proteins of milk. The proteins of the inner endosperm constitute 75% of the total grain proteins but they contain a less well-balanced and incomplete mixture of amino-acids. The *mixed* proteins of the *whole* grain have a high biological value as the main endosperm deficiencies are compensated for by the proteins of bran and germ.

(iv) The minerals (Ca, P), the vitamin-*B* complex (aneurin, riboflavin, nicotinic acid) and vitamin-*E* are mainly concentrated in the bran and germ, as is the phytic acid (see Fig. 656).

It is now possible to compare the relative merits of different extraction rates. Eighty-five per cent. flour compared with 70% flour is slightly darker, keeps less well, and goes mouldy more readily; its phytic acid content is higher leading to formation of insoluble Ca phytate and loss of food Ca; its fibre content is higher; these are the disadvantages. On the other hand 85% flour has a higher content of proteins of superior biological value and its content of Ca, Fe, vitamin-*B* complex, vitamin-*E* (and possibly other unknown trace substances) is higher; to compensate for its phytic acid it contains more of the enzyme *phytase* which hydrolyzes phytic acid while the dough is being made and rising; the harmful effect of the phytic acid can be safely annulled by fortifying the flour with added Ca in the form of $CaCO_3$; the colour is a matter of conditioning (people can be persuaded

to like what they get). British experts in nutrition have strongly advised that the standard flour produced should be 85% extraction and that millers should not be allowed to revert to the physiologically inferior pre-war 70% flour.[1]

Nutritional Needs of Pregnancy and Lactation.—The body of a baby (weighing 7 lbs.) at birth contains about 500 g. of protein, 30 g. of Ca, 14 g. of P, and 0·4 of Fe; over two-thirds of this is laid down in the last 3 months and one-third in the last month of pregnancy. In addition the mother lays down new protein in the growing uterus, breasts, and other tissues (perhaps another 500 g.). The basal metabolism of the mother at the end of pregnancy has increased by 25% or by about 350 Cal. daily. After birth the baby doubles in weight from 7 to 14 lb. at 3 months and increases to 21 lb. at 1 year; each 7lb. represents an amount of new tissue equal to that formed throughout pregnancy. 100 c.c. of human milk contain 1·2 g. of protein and 0·03 g. of calcium, and yield 54 Cal. At one month after delivery the mother must supply the baby with 7 g. of protein, 0·18 g. of Ca and 300 Cal. daily; at 6 months these amounts have risen to 9·5 g. of protein, 0·25 g. of Ca, and 450 Cal. daily. The milk proteins (lactalbumin and caseinogen) have a specific amino-acid constitution and can only be formed if all the necessary amino-acids are supplied in the food in sufficient amounts; the same applies to the new tissues formed by mother and child during pregnancy. It is clear that the pregnant and lactating woman

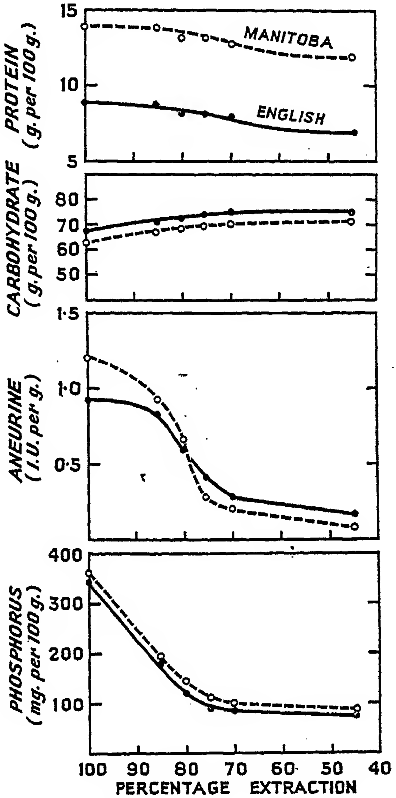


FIG. 656.—Effect of Rate of Extraction on Composition of Flour. Canadian and English Flour Compared. (McCance, *Lancet*, 1946, i, 79.)

[1] AGENE POISONING.—Flour is commonly bleached by millers with NCl_3 (the agene process) to improve its colour and baking qualities. When biscuits made from such flour forms the main constituent of the diet of dogs they develop "canine hysteria": in a severe attack the animal sits first in a sphinx-like attitude; its head then starts jerking and the movements spread throughout the body. Next it starts running round its cage, sometimes barking furiously and either dashes into the walls without apparently seeing them or attempts to jump up them. It may then stop running and stagger round as though in a drunken state and gradually recover or else develop an epileptiform fit. Experiments have shown that agene interacts with the proteins of flour (gliaden and glutelin) to form a toxic substance which is derived from methionine.

The use of agene has been banned in the U.S.A. and Canada. Its place as an "improver" is being taken by chlorine dioxide, which in concentrations far greater than those used commercially produces no harmful effects in dogs or other animals. (Mellanby, *Brit. med. J.*, 1946, ii, 885; 1947, ii, 288.)

needs a *substantially increased protein intake*, a considerable part of which should be of animal origin. The recommendations are: for the last 3 months of pregnancy a minimum of 1·5 g. of protein per kg. (*i.e.* an extra 30 g.) daily; during lactation the allowance should be higher, up to 2 g. of protein per kg. The total caloric requirements are also increased to the extent indicated.

Calcium requirements are discussed on p. 1000, vitamin-*D* on p. 1010, and iron on p. 205.

The vitamin content of mother's milk (and cow's milk) depends on the maternal diet. A good specimen of human milk contains, per 100 c.c.: vitamin-*A*, 300 i.u.; thiamine, 10 i.u.; -*C*, up to 6 mg. (three times that of cow's milk); and -*D*, 10 i.u. It is advisable to supplement both the vitamin -*C* and -*D* intake of the baby and the -*D* intake of the mother (p. 1064).

Great stress must be placed on the importance of *milk as a "protective food" especially for children and expectant or nursing mothers*, who should drink at least 1 pint daily. It is the only naturally balanced food containing nearly 20 g. of first-class proteins per pint, minerals (especially calcium and phosphate), all the vitamins (especially -*A* and -*D*), and fat.

Mellanby[1] has made the following recommendations for a diet during *pregnancy*: Milk, 1–2 pints; green vegetables and an egg once or twice daily; sea fish (for iodine content) every other day; calf's liver once weekly. If the diet is deficient it should be supplemented with 2 ounces of cod-liver oil daily. The rest of the diet can be within reason whatever the woman likes. It should never be forgotten that a pregnant woman has to provide in her diet for the *full requirements of the growing fœtus. The future of the child depends to a great extent on the adequacy or otherwise of the maternal diet during pregnancy and lactation.*

British Rationing System.[2]—The British food rationing system which was operated during and after the last war is a brilliant example of the application of our knowledge of the physiology of nutrition to the interests of a community faced with food scarcity. Underlying the rationing system there was also a moral principle which might well be adopted by all civilized peoples: that every member of the community must have enough before anyone may have more. The rationing system ensured that in spite of war-time scarcities the underprivileged 20% of the British people who before the war did not get enough to eat were also properly fed. Particular attention was paid to the requirements of special classes like children, pregnant and lactating women, and heavy manual labourers. It is the view of the British Ministry of Health that "the national provision of milk and vitamin supplements to the priority groups has probably done more than any other single factor to promote the health of expectant mothers and young children during the war." Between 1938 and 1947 the infant mortality rate per 1000 live births fell in England and Wales from 53 to 41 and in Scotland from 70 to 56; the corresponding figures for 1950 were 30 and 39.

The rationing system as it was enforced towards the end of the war is described below in some detail. After the war, as food became more plentiful, an increasing number of foodstuffs were freed from control. In these uncertain times one never knows when the rationing system may have to be extended once more.

[1] The advice was given before the outbreak of the second world war; the advice is still good if not always or everywhere practicable.

[2] Drummond, *Nutritional Requirements of Man in Light of War-time Experience*, 1947 (Roy. Inst. Chem. Lecture).

In Britain the foodstuffs were divided into: (i) rationed (the maximum amounts of which were fixed; (ii) regulated (controlled by more flexible orders); (iii) pointed (to each of a number of foods was assigned a point value, and the total number of points per head fixed for a given period); (iv) price fixed (as were (i), (ii), and (iii)); and (v) uncontrolled (bought by arrangement between buyer and seller). A child over 5 years counted as an adult; younger children got smaller rations but were granted other privileges. Food could, in addition, be obtained in schools, canteens, and restaurants without surrender of coupons.

Rationed and Regulated Foods.—The table on p. 1062 sets out full details of the 1944 arrangements for rationed and regulated foods. They provided daily about 30 g. of animal protein (which is probably an adequate total intake of first-class protein of high biological value), 60 g. of fat, half the necessary calcium and vitamin-*A*, and yielded nearly 1000 Cal.; these foods are however almost devoid of iron and vitamin-*C*, and are very inadequate in thiamine and riboflavin content. The price of the rationed foods was kept very low by means of heavy subsidies.

Food on Points.—The point system is an ingenious device for controlling the distribution of certain classes of food which are available in insufficient amounts. The classes of food so controlled were: (i) most canned fish and all canned meat; (ii) dried and canned peas; beans and lentils; (iii) sweet biscuits; (vi) rice, sago, and tapioca; (vii) syrup; (viii) canned fruit. A certain number of points were allotted per head per month. The pointed foods enabled certain inadequacies in the diet to be dealt with; they provided a choice of animal proteins, concentrated vegetable proteins, certain vitamins and, what is perhaps equally important, they permitted of some variety in the diet. Each person could exercise his own discretion in the use of his points. By raising or lowering the point value of any article of food the demand could be adjusted to the supply available at the time without manipulating prices and so hitting the poorer consumer. An unpopular food could be kept at a low point value initially to encourage trial usage; when a public taste for it had been established the point value could be raised.

Supply of Calories.—About 1000 Cal. daily were supplied by the rationed foods. The consumption of these and certain other foods was of necessity greatly reduced below pre-war levels.

The rationed food and the foods obtained on points provided in all about 1300 Cal., leaving 1700 Cal. still to be found to make up the necessary 3000 Cal. daily; these calories came predominantly from bread and flour and potatoes.

The pre-war consumption of bread and flour was 66 oz. per head per week; during the war it rose to 72 oz. or more. Pre-war bread and flour supplied about 800 Cal. per day; it now provided about 900 Cal. The pre-war average potato consumption was 9 oz. per head per day, yielding 210 Cal.; the war-time consumption rose by one-third to 12 oz., yielding 280 Cal.

The total daily calorie intake from the above-mentioned sources was about 2300 Cal.

Rationed food . . .	950 Cal.
Pointed foods . . .	200 ,,
Bread and flour . . .	900 ,,
Potatoes	280 ,,
Total .	2330 Cal.

RATIONED AND REGULATED FOODS (VALUES PER WEEK), 1944

FOOD.	Amount.	Calories.	g. Protein.	g. Fat.	g. Carbo-hydrate.	g. Calcium.	mg. Iron.	i.u. -*A*.	i.u. Thia-mine.	i.u. Riboflavin.	mg. -*C*.	i.u. -*D*.
Milk . .	2 pints	800	36	44	52	1·36	0·8	1,650–6,600	250	300–900	20–50	20–200
Cheese . .	3 oz.	510	21	24	..	0·60	0·3	variable + +	..	..	..	..
Sugar . .	8 oz.	930	..	..	225	..	..	..	..	..	..	..
Butter . .	2 oz.	450	..	48	..	..	..	1,000–3,000	20	..	..	12–90
Margarine .	4 oz.	900	..	97	..	..	..	2,200	..	..	..	240
Cooking fat .	2 oz.	440	..	46	..	..	..	..	..	..	..	..
Bacon . .	4 oz.	480	32	40	..	..	4·4	..	90	30–90	..	..
Tea . .	2 oz.	..	..	..	..	..	..	..	..	..	..	..
Meat . .	12 oz.	av. 960	av. 90	av. 60	..	..	15	..	500	..	..	..
Preserves . .	4 oz.	325	2	..	80	..	0·8	..	..	..	..	..
Household milk .	2 oz.=1 pint	190	20	..	26	0·7	0·6	..	80	..	..	..
Dried eggs .	1½ oz.=3 eggs	115	16	11	..	0·8	5	1,000	120	480	..	30
Shell egg . .	½ egg	46	4	4	..	..	0·7	250–1,100	10	..	..	40
Chocolate . .	3 oz.	480	5	20	45	..	2·1	..	..	..	..	..
TOTALS .	..	6,630	226	394	428	3·46	29·7	6,100–13,900	1,070	810–1,470	20–50	342–600
	Per day .	950	32	56	61	0·43	4·2	870–1,990	150	115–210	3–7	49–86

Notes: The meat ration was fixed by value; it purchased ½ lb. to 1 lb. of meat depending on cut and quality; average values are given.
Margarine is fortified with the amounts of vitamins-*A* and -*D* indicated.
The vitamin-*A* and -*D* contents of milk, cheese, and butter vary widely with the season and the diet of the cow.
The milk ration indicated was the minimum winter ration; more milk was available in the summer, up to 3 or 4 pints per week.
Additional amounts of household milk and dried egg were often available.
Margarine is fortified per oz. with -*A*, 550 i.u.; -*D*, 60 i.u.
Composition of dried egg/100 g.: protein, 47 g.; carbohydrate, 1 g.; fat, 31 g.; calories, 330; -*A*, 2,860 i.u.; Thiamine, 360 i.u.; -*D*, 100 i.u.
Dried milk per oz.: protein, 10 g.; carbohydrate, 13 g.; fat trace; calories, 94 Fe, 0·3 mg. Thiamine, 37 i.u.; no -*A* or -*D*.

As no allowance has been made above for miscellaneous purchases of unrationed food, vegetables, or extra milk, it is probable that the requirements of the sedentary worker were reasonably well covered. Heavy manual workers had to eat larger amounts of bread and potatoes and were in difficulties as they were unable to get the extra fat upon which they normally relied. Canteens were established in connection with most factories, which received additional supplies of rationed food; special preference was given to heavy workers who were thus given significantly larger rations. Cheap eating houses, so-called British Restaurants, and expensive hotels, enabled those who had the money, time, or opportunity to obtain meals away from home, and enabled large additional sections of the population to supplement their diet to a variable extent. The housewife who had to get all her meals at home and often sacrificed her own rations to feed the rest of her family and who was often the hardest worked and most worried and frustrated member of the family was among the principal sufferers.

PROTEINS.—The rations (per day) supplied 30 g. of animal protein; bread and potatoes provided about 30 g. of vegetable protein. There was no difficulty in attaining or exceeding the 70 g. minimum.

FAT.—The rations supplied 60 g. daily, so minimal needs were fully covered.

CARBOHYDRATE.—The intake need only be considered from the standpoint of total calories.

VITAMINS AND MINERALS.—The key substances outside the rationed and pointed foods were bread, green vegetables, carrots, and potatoes.

Bread.—As explained on p. 1058, 85% extraction flour was used with its better content of proteins, iron and the vitamin-*B* group. Its calcium was less available for absorption owing to the larger *phytic acid* content of the flour which leads to precipitation of insoluble non-absorbable Ca phytate; but as the flour was fortified with $CaCO_3$ (60 g. of Ca per 100 g. of flour) (which spoils neither the taste nor the colour) the available Ca of "wheatmeal" was increased to 80 mg. per 100 g.

Vegetables.—The consumption of green vegetables, potatoes and carrots increased greatly, providing additional supplies of vitamin-*A* and -*C*.

Vitamin-A.—1000–2000 i.u. daily were supplied by the rations. 250 g. of green vegetables give 2500 i.u. and 10 g. of carrots another 400 i.u. The standard 4000 i.u. was easily covered.

Thiamine.—The rations supplied daily 150 i.u.; 10 oz. daily of wheatmeal bread (say 300 g.) contain 150 i.u.; potatoes and vegetables at pre-war level yield 120 i.u. The minimum of 400 i.u. was thus easily exceeded.

Rest of -B Complex.—The low meat ration greatly reduced the riboflavin and nicotinic acid intake; this was compensated to some extent by the wheatmeal bread; there was also a gain in riboflavin from the raised consumption of vegetables.

Vitamin-C.—The rationed foods were devoid of -*C*. Pre-war, the average intake was 60 mg. daily provided mainly in equal thirds by potatoes, fruit, and green vegetables; the absence of fruit was compensated for to some extent by the increased intake of green vegetables. Probably 30 mg. of vitamin-*C* daily are adequate for normal health (p. 1042).

Vitamin-D.—Some additional -*D* was provided by the pointed (and also the uncontrolled) fat fish; *e.g.* 2 lb. of herring per month (=12 points) gave an

extra 180 i.u. of *-D* per day and thus made up (with 80 i.u. in the rationed food) the *-D* needs.

Calcium.—The rationed foods provided 0·5 g.; 300 g. of "wheatmeal" contain 0·25 g. of available calcium; the minimal needs were thus covered.

Iron.—The iron intake was possibly deficient for women during the reproductive period; any shortage could easily and cheaply be made good by taking medicinal iron salts.[1]

Rations for Special Groups.—"Juniors" (aged 5–16 years) received the full adult rations and a priority supply of milk ($3\frac{1}{2}$ pints weekly). As about $\frac{1}{3}$ pint of milk daily was also supplied free at school, the total milk available was 5 pints weekly. Good cheap dinners were and are provided in an increasing number of schools. Children under 5 years of age also received adult rations (only a half-ration of meat and no tea), 7 pints of milk weekly at less than one-third the current prices, extra eggs, and extra orange juice and cod-liver oil. Pregnant and lactating women received their adult ration and *in addition the rations allotted to a young child.* They thus received, generally speaking, double rations (except meat and tea), 7 pints of milk, and some eggs above the adult ration. Pregnant and lactating women also received extra orange juice and vitamins-*A* and *-D* in cod-liver oil or in tablet form (each tablet contained 4000 i.u. of vitamin-*A*, 800 i.u. of vitamin-*D*, 250 mg. of calcium phosphate and 0·13 g. of potassium iodide).

[1] Suggestions were made during the last war to add iron salts to the flour, which was already being fortified with chalk. But fears that the resulting bread might be called ferro-concrete prevented the adoption of this otherwise (nutritionally) admirable proposal.

XI

REPRODUCTION AND ITS ENDOCRINE CONTROL

STRUCTURE AND FUNCTIONS OF THE OVARY[1]

The Ovarian Cycle (Fig. 657).—(1) Primordial Follicles.—During fœtal life the surface of the ovary is covered by a layer of small cubical cells, the *germinal epithelium.* These cells multiply and grow down into the substance of the ovary; some cells enlarge and differentiate to become ova, and become surrounded by a single layer of cells forming the *membrana granulosa*; each ovum, plus its surrounding membrana granulosa, is called a primordial follicle. It is asserted that the ovary of a newborn baby girl may contain 30,000–300,000 ova. It is uncertain whether fresh ova are budded off from the surface epithelium after birth. Before puberty many of the primordial follicles enlarge because of proliferation of the surrounding granulosa cells; but the whole structure subsequently degenerates, dies, and is absorbed by phagocytic cells. (This process is called *follicular atresia*).

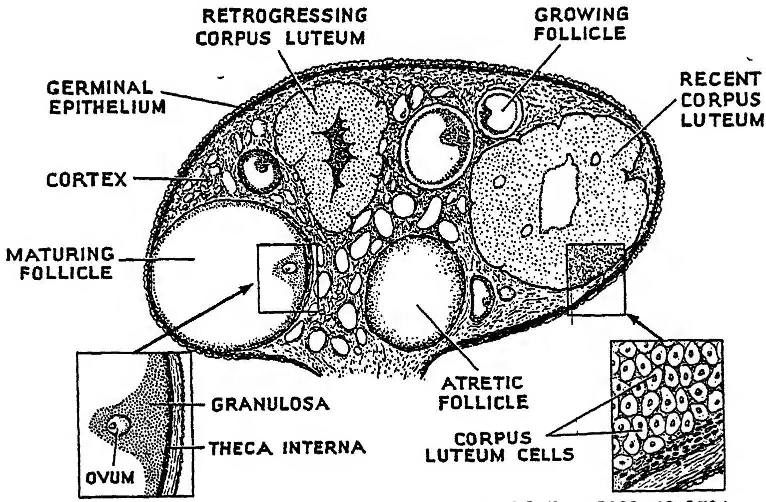


Fig. 657.—Structure of the Ovary. (Corner, *Physiol. Rev.*, 1938, *18*, 156.)

[1] Marshall, *Physiology of Reproduction,* London, 2nd edn., 1922. Parkes, *Internal Secretions of Ovary,* London, 1929. Robson, *Recent Advances in Sex Physiology,* London, 3rd edn., 1947. Allen (editor), *Sex and Internal Secretions,* Baltimore, 2nd edn., 1939. Pincus and Thimann, *The Hormones,* N.Y., 1948, *1*; 1950, *2*. Symposium, "Physiological and Psychological Factors in Sex Behaviour," *Ann. N.Y. Acad. Sci.*, 1947, *47*. 603–664.

At puberty the ovary contains tens or hundreds of thousands of these immature primordial follicles; of these only a few hundred will mature and give rise to the ova which are discharged from the ovary monthly during thirty or more years of active reproductive life; there has thus been a vast overproduction of ova in relation to adult needs.

(2) Graafian Follicles.—At puberty full follicular growth sets in. Towards the end of a menstrual period one primordial follicle rapidly develops. The ovum enlarges and becomes surrounded by a thick tough membrane, the *zona pellucida*; the granulosa proliferates to form many layers of small cells. Fluid then appears in these cells, first as droplets, which coalesce to form a cavity containing the *liquor folliculi*. The liquor separates the granulosa into two layers: (i) a thin layer which lines the inner surface of the follicle; (ii) cells which surround the ovum and which form a mass (*cumulus, discus proligerus*) at one pole thus attaching the ovum to the margin of the follicle. The mesenchyme external to the Graafian follicle differentiates to form a thin *theca interna* (of small deeply staining cells) supported by an outer vascular layer.

(3) Ovulation.—More fluid exudes into the follicle, and the pressure within it rises. The ovarian substance which separates it from the surface necroses and gives way, and the follicle ruptures. This process is known as *ovulation*. The ovum is loosened, floats freely in the liquor, escapes into the peritoneum and enters the Fallopian tube. Ovulation usually occurs between the thirteenth to seventeenth day after the first day of the menstrual period in women but may occur earlier or later. This has been shown by direct inspection of the ovaries at operation and by the recovery of ova from the Fallopian tubes. By the time of ovulation the ovum has undergone its *reduction division* which halves its chromosome number.

Time of Ovulation.—An accurate knowledge of the time of ovulation in women is not of academic interest only. It is not known for certain how long an ovum survives after discharge from the ovary or how long the sperms can live after being introduced into the vagina; the evidence suggests that in neither case are they functionally active after an interval of several days. For pregnancy to occur it is therefore necessary that coitus should take place within a few days either side of ovulation, which is presumably the time of maximum fertility. Coitus at other times would tend to be sterile, and it is suggested that the rest of the menstrual cycle constitutes a more or less "safe period," *i.e.* pregnancy is unlikely to occur even if no other methods of "birth control" are employed. This method will, however, fail in its purpose if ovulation in any month is premature or delayed, and such variations occur frequently. There is evidence that ovulation may occur even during the latter part of the menstrual period.

Fallopian Tube.—The ovum is wafted along the Fallopian tube towards the uterine cavity partly by the lining ciliated epithelium (which is tallest at the time of ovulation) and partly by contraction of the muscular wall. This transfer of the ovum from ovary to tube is much facilitated by the slow sweeping movements of the infundibulum over the surface of the ovary; these are greatest in this phase of the cycle. The activity of the tube is influenced by hormones; its motility is increased by œstradiol and depressed by progesterone.

(4) Corpus Luteum.—At ovulation the lining membrane of the follicle

collapses into a folded layer. The blood vessels of the surrounding vascular layer leak and blood extravasates into the cavity of the follicle and coagulates, forming the *corpus hæmorrhagicum.* The membrana granulosa proliferates and differentiates to give rise to large pointed epithelial cells containing large nuclei and yellow pigment (lutein). These cells, accompanied by small blood vessels and thecal connective tissue elements grow into and replace the blood clot, forming the *corpus luteum.* It matures at about the nineteenth day, and if pregnancy does not occur it persists till just before the onset of the next menstrual period, when it begins to degenerate. Degeneration is completed during the menstrual period: the connective tissue cells form collagen fibres which obliterate the blood vessels; the corpus luteum loses its pigment, shrinks, and is finally replaced by an avascular scar. The duration of activity of the corpus luteum of menstruation is fairly constantly 12–14 days, unlike the period of follicular growth and ovulation which is extremely variable. If pregnancy occurs the corpus luteum continues to grow for several months, attains a large size and begins to degenerate at about the sixth month.

The control of the ovarian cycle by the anterior pituitary *gonadotrophins, i.e.* follicle-stimulating hormone (FSH), luteinizing hormone (LH), and prolactin (luteotrophin) is considered on p. 1083.

INTERSTITIAL CELLS.—There are none in the adult human ovary. When well developed, *e.g.* in the rabbit, they consist of typical epithelial cells, without ducts, and with an abundant blood supply. The cells are polyhedral in shape, and contain numerous granules, chiefly of a lipide nature, which may represent their secretory product. They are not found at all in about half the species which have been examined. Little is known about their functions (cf. p. 1104).

The changes in the ovary at the *menopause* are described on p. 1069.

Functions of the Ovary.—In addition to its essential function of discharging *ova* (p. 1066), the ovary secretes two hormones: (i) an *œstrus-producing* (*œstrogenic*) substance called *œstradiol*, formed perhaps by the follicular tissue; (ii) *progesterone* formed by the corpus luteum. By means of these two hormones the ovary regulates the activities and nutrition of the rest of the reproductive organs. It is responsible for the following changes:

(i) The growth and development of the *uterus*, *Fallopian tubes*, and *vagina* at puberty.

(ii) The *œstrus cycle* in lower mammals (p. 1074) and the changes of the *menstrual cycle* in women (and apes and monkeys).

(iii) The appearance and, in some cases, the persistence of the *secondary sexual characters*.

(iv) Certain of the bodily changes that take place during *pregnancy*, especially the *embedding of the ovum* in the uterus and the development of the *placenta*.

(v) Some of the changes which take place in the *mammary glands* (p. 1092).

Puberty.—This is the period during which the ovary and the accessory reproductive organs (*e.g.* uterus, vagina, breasts) begin to grow and develop, and secondary sexual characters begin to appear These processes begin at the age of 10–14 years. Puberty is regarded as ending with the occurrence of the first menstrual period, usually at 13–15 years, the range being 10–15

years. The main changes that take place at puberty are summarized below.

(i) Complete ovarian cycles occur characterized by ovulation and corpus luteum formation. The onset and recurrence of the ovarian cycle is due to the action of the gonadotrophic hormones of the anterior pituitary which begin to be secreted at this time. The factors initiating anterior pituitary activity are unknown.

(ii) The uterus and vagina enlarge ; the growth curve of the reproductive organs is shown in Fig. 600, D. The muscle fibres of the uterus increase in number and size ; the mucous membrane thickens and the gland alveoli become larger. The pubertal changes in the uterus and vagina can be produced experimentally in immature animals by injection of œstrogens.

(iii) The breasts begin to appear as a result of outgrowths of ducts from the nipple area, and an increase in the amount of fat, connective tissue and blood vessels. Such changes as occur before the onset of ovulation are due entirely to œstrogen; during adolescence, owing to the action of both œstrogen and progestin, glandular alveoli appear and progressively increase in size and numbers with each ovarian cycle (cf. p. 1092).

(iv) The secondary sexual characters develop ; these include the female distribution of fat, giving the characteristic curves to the body, and the appearance of hair in the axilla and on the pubes (in the latter region the upper hair margin is concave upwards).

(v) Important psychological changes take place as the girl matures mentally and emotionally through adolescence to young womanhood.

Extirpation of Ovary.—(1) Before Puberty.—Little is known about the results of extirpation of the ovary before puberty in girls. Probably puberty does not set in, the menstrual flow does not appear, and the secondary sex characters do not develop. It is obvious that the presence of the ovary is essential for the onset of puberty, though *indirectly* the anterior pituitary is the dominant factor (p. 1083).

(2) In Adults.—Following extirpation of the ovary in adults, there is atrophy of the whole genital apparatus—the uterus, the vagina, and the external genital structures. Menstruation ceases permanently. Vasomotor changes are common, *e.g.* flushing of the skin of the face, neck, and upper chest (" hot flushes "), sweating and a feeling of suffocation. The effects on the breasts are variable : they may increase in size, owing to local accumulation of fat, or they may shrink because the glandular tissue atrophies. Obesity develops from diffuse deposition of fat. Conflicting reports are given concerning the effect on sexual desire, but it is often unaffected ; thus in women as in men, though sexual desire is modified by sex hormones it may in large measure be independent of them and be determined by nervous (*i.e.* emotional and mental) factors. It is also quite certain that sexual desire may persist, sometimes to a heightened degree, in women after ovarian atrophy at the menopause. Complete ovariectomy may result in considerable emotional disturbance, varying from a certain amount of irritability or depression to a condition closely allied to insanity (cf. the results of castration in the male, p. 1108).

Menopause.—The menopause (climacteric) is the period of life when menstruation naturally ceases and other phenomena identical with those just described make their appearance. It usually occurs between the ages of forty-five and fifty, although it may set in earlier or later. The condition

is associated with marked changes in the ovaries: they become smaller, the Graafian follicles disappear and are replaced by fibrous tissue; ova, corpora lutea, and the internal secretions of the ovary are no longer formed. These ovarian changes are not due to lack of anterior pituitary hormones, but to a "senile" change [1] in the ovary itself which no longer reacts to the hormones which normally stimulate it.

Ovarian Grafts.—When the ovaries are grafted into ovariectomized women, they may function for a time, usually for about twelve months. During this period they remain normal in structure, and undergo cyclic ovulation. Degeneration then sets in, affecting first the corpora lutea and later the follicles which become cystic. While the grafts function, normal cyclical changes take place in the other organs of reproduction (*e.g.* the uterus). These experiments prove conclusively that the ovary forms one or more internal secretions.

THE HUMAN MENSTRUAL CYCLE

Structure of Uterus.—The body of the uterus consists of the following layers from without inwards:

(i) Serous coat.

(ii) Thick muscular coat supplied by sympathetic nerve fibres.

(iii) Mucous membrane or endometrium which undergoes characteristic changes during the menstrual cycle.

Two kinds of blood vessels enter the endometrium from the deeper layers:

(i) *Spiral* arterioles which take a very tortuous course to the surface to break up into capillaries which supply the superficial third of the mucosa.

(ii) *Straight* arterioles which run for only a short distance to supply the basal two-thirds of the mucosa (Fig. 658).

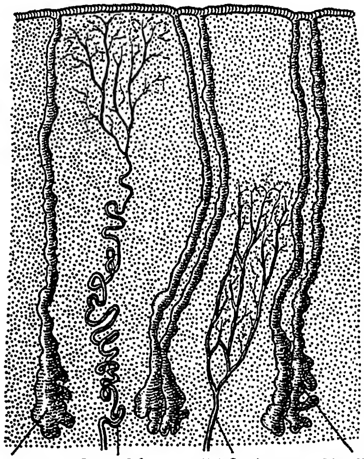


FIG. 658.—Blood Vessels of Uterine Mucous Membrane. (After Corner, *Hormones in Human Reproduction*, Princeton, 1946.)

NOTE.—(i) Uterine glands; (ii) spiral arteries supplying inner third of mucosa; (iii) straight (basal) arteries supplying basal part of mucosa.

Stages of Menstrual Cycle [2] (Figs. 659, 660).—(1) PROLIFERATIVE STAGE.—When the damage resulting from the menstrual period has been fully repaired (*e.g.* on about the fifth or sixth day) the proliferative phase begins. The endometrium is initially thin (*e.g.* under 2 mm.) and

[1] The term "senile" is hardly a happy one; the ovarian change is one which normally occurs at a particular age.

[2] It should be emphasized that the menstrual cycle is unrelated to the œstrous cycle of lower mammals (*e.g.* rodents).

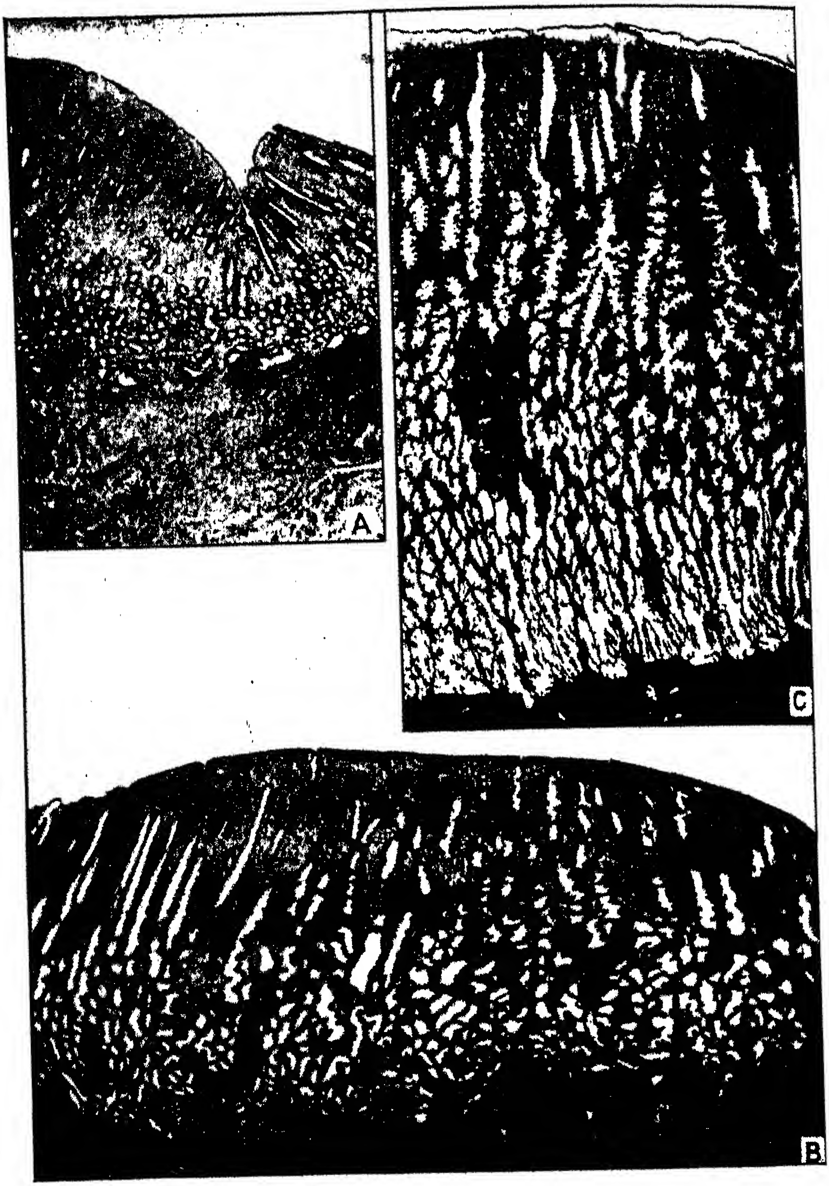


FIG. 659.—Endometrial Changes during the Menstrual Cycle (Rhesus monkey). (Preparations by Corner, Hartman, and Bartelmez, from Corner, *Hormones in Human Reproduction,* Princeton, 1946.)

A. Fifteenth day of cycle, just after ovulation (proliferative phase).
B. Twenty-third day of cycle, progestational phase.
C. Twenty-seventh day of cycle; full progestational changes; menstruation due one day later.
All sections are magnified by x10.

consists of a ciliated columnar epithelium (which beats towards the exterior), dipping down into a loose stroma to form simple tubular glands. During the next 8 days or so (*i.e.* 6th–14th day) the mucosa thickens, becomes more vascular, and the glands elongate and become dilated in their deeper part.

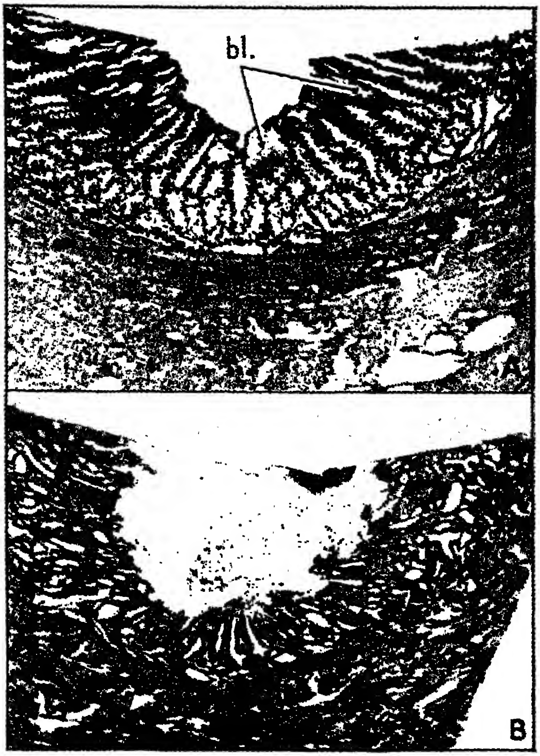


FIG. 660.—Endometrial Changes during Menstruation (Rhesus Monkey). (Corner, *Hormones in Human Reproduction*, Princeton, 1946.)

A. First day of flow; *bl.*, small collections of blood in the lining of the uterus. Progestational pattern of glands still present.

B. Third day of flow. Note loss of superficial part of endometrium; the progestational pattern of the glands has disappeared.

(2) PREMENSTRUAL (PROGESTATIONAL) STAGE (15th–28th day).—The endometrium progressively increases in thickness, *e.g.* to 4–5 mm. The outstanding features are the increase in the length and diameter and the change in the outline of the glands; they are greatly distended with mucus and the lining is thrown into folds which project into the lumen giving the gland wall a saw-edge, tufted appearance. The stroma cells proliferate and enlarge, become more closely packed and resemble those seen in the early placenta. The blood vessels are congested and exudation of clear and blood-stained fluid occurs.

(3) MENSTRUATION (1st–4th or 5th day).—This stage is characterized by bleeding and shedding of the superficial part of the endometrium, leaving the basal (deeper) layer intact. The mechanism is obscure. It is suggested that the spiral arteries close down (perhaps for hours). Owing to the resulting ischæmia the related region of the mucosa undergoes necrosis and the walls of the contained capillaries are weakened. When the spasm passes away and the circulation is restored, blood leaks out through the damaged and destroyed areas of the capillary wall into the stroma, under the superficial epithelium, and into the lumina of the glands. The necrotic endometrium, together with exuded blood and much mucus, is cast off into the lumen of the uterus, whence it passes to the exterior. Menstrual blood which escapes rapidly from the uterus clots promptly, yields abundant fibrin, and contains thrombin. If the blood is retained longer in the body of the uterus, partial clotting takes place there or films of fibrin may be deposited on the endometrium. Intra-uterine clots when long retained undergo a secondary process of liquefaction.

It must not be supposed that this destructive process occurs simultaneously throughout the endometrium; if it did the whole of the superficial part of the mucosa would be rapidly sloughed off and a sudden large hæmorrhage would occur. At any one time the process is probably affecting only small scattered areas of the endometrium. The flow of blood into the vagina takes place in tiny spurts or trickles at intervals of 1–15 minutes. By the end of 4 or 5 days the whole endometrium has been affected, the destructive phase in any area being soon followed by repair. The surface epithelium and the simple tubular glands are restored by outgrowths from the persistent basal ends of the glands.

The total blood loss in normal women in one menstrual period varies from 10–210 c.c. with a mean of 40 c.c. There are variations in the length of the complete menstrual cycle in any one woman; no woman is absolutely regular and the range of cycle length may be 28 ± 9 days.

ANOVULATORY MENSTRUAL CYCLE.—In rhesus monkeys anovulatory cycles may occur in which no ovulation takes place and no corpus luteum is formed. The endometrium then only shows the changes characteristic of the proliferative stage; the duration of these cycles is the same as that of a normal ovulatory cycle. When bleeding occurs it is similar in duration and is accompanied by the same kind of vascular changes as in ovulatory cycles. Many workers urge that the bleeding at the end of an anovulatory cycle should also be designated menstruation. Anovulatory cycles also occur infrequently in women, chiefly in young girls and near the time of the menopause.

Mechanism of Menstrual Cycle.—(1) EXPERIMENTAL UTERINE BLEEDING (Fig. 661, A).—(i) When both ovaries are removed in monkeys (or women), a single episode of bleeding occurs after a few days, irrespective of the phase of the cycle at which the operation is performed. The bleeding is thus related to *withdrawal* of an ovarian factor (or factors) which influences the uterus.

(ii) If a monkey is ovariectomized and immediately afterwards given a course of œstrogen injections, no bleeding occurs after the operation. But as soon as the œstrogen treatment is stopped bleeding occurs from the endometrium which is in the proliferative phase (Fig. 661, B). This experiment seems to reproduce what takes place in anovulatory cycles in intact

animals in which only œstrogen is secreted; in such cycles, when the unruptured follicle begins to degenerate and stops secreting œstrogen, bleeding takes place. Bleeding can thus result from *œstrogen deprivation* (or even from a lowering of the previous blood œstrogen level).

(iii) An ovariectomized monkey is given œstrogen to build up the proliferative phase of the endometrium. Œstrogen treatment is then stopped and is replaced by injections of progesterone. No bleeding occurs in spite of œstrogen deprivation; the presence of *progesterone inhibits the bleeding* and

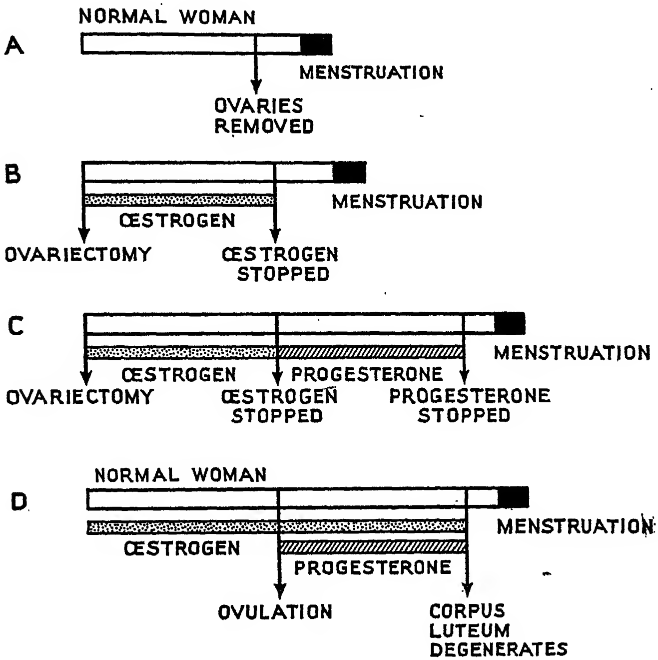
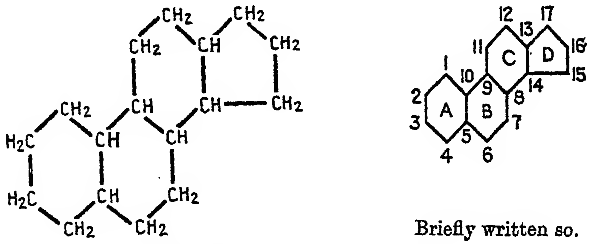


FIG. 661.—Diagram to Illustrate Hormonal Control of Menstrual Cycle.

also builds up the progestational phase of the endometrium. On stopping the progesterone treatment, the endometrium which has been deprived of the influence of both œstrogen and progesterone bleeds (Fig. 661, C).

(2) MECHANISM OF NORMAL MENSTRUAL CYCLE (Fig. 661, D).—In the normal menstrual cycle the sequence of events is probably as follows. During the growth of the Graafian follicle up to the stage of ovulation, œstrogen alone is secreted by the ovary and gives rise to the proliferative phase. When the corpus luteum develops it secretes progesterone and also continues to form œstrogen. Under the influence of the combined action of these two hormones the progestational stage is produced. When the corpus luteum degenerates the endometrium is deprived simultaneously of the "supporting action" of both hormones, and bleeding occurs. It must be presumed that the hormones have a specific action on the spiral arteries of the endometrium.

In its spatial arrangement the ring system is relatively flat (*i.e.* in the plane of the paper, as conventionally written). Substituents (replacing any H) stick out on one side or other of the plane of the molecule. Those

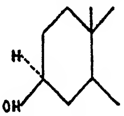


Briefly written so.

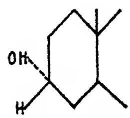
FIG. 662.—Structure of Steroids.

projecting towards the observer are termed β substituents, and those projecting away are called α substituents.

The direction of the projection is shown thus :



3-β hydroxy compound.



3-α hydroxy compound.

The *continuous* line indicates projection *towards*, and the *broken* line projection *away from*, the observer.

In naming the steroids the following conventions are used :

-ane: fully saturated compound.

-ene: unsaturated compound ; *-diene, -triene,* represents the presence of 2 or 3 double bonds.

-one: presence of ketone (: C=O) grouping.

-ol: presence of hydroxyl (alcoholic or phenolic —OH grouping).

PHYSIOLOGICAL STEROIDS.—The steroids of physiological interest include :

(i) Œstrogens (*infra*).

(ii) Progestins (p. 1077).

(iii) Androgens (p. 1076).

(iv) Adrenal cortex hormones (corticoids) (p. 959).

(v) Cholesterol and related substances like coprosterol and ergosterol (p. 799).

(vi) Vitamin-*D* precursors (p. 1010).

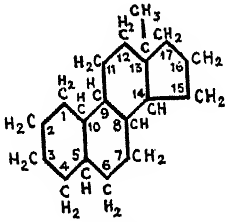
(vii) Bile acids (*e.g.* cholic acid) (p. 798).

(viii) Cardiac glucosides like digitoxigenin (of digitalis).

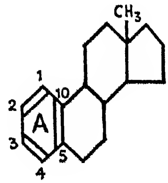
(ix) Carcinogens, like methyl cholanthrene.

ŒSTROGENS (FOLLICULOIDS) (Fig. 663).—The structure is based on *œstrane*, which differs from the basic steroid shown in Fig. 662 in having a β CH_3 group replacing the H attached to the C in the 13 position. There are no double bonds ; all the C atoms have enough H atoms to fill up their quota of 4 valencies.

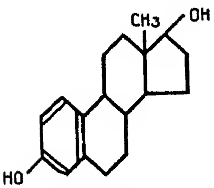
A derivative of œstrane is *œstratriene* (three double bonds; these join the following carbons: 1 : 2; 3 : 4; and 5 : 10, to give an aromatic ring, A).



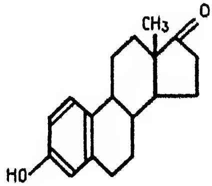
Œstrane.



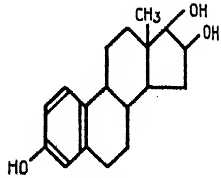
Œstratriene.



Œstradiol
In ovary. Activity: 6 × œstrone.



Œstrone
In urine.



Œstriol.
In urine; weaker than œstrone.

FIG. 663.—Structure of Œstrogens.

The natural œstrogens are substituted derivatives of œstratriene, thus:

(i) *Œstradiol* (-diol=dihydroxy): the two OH groups are in positions 3 and 17. The 3-OH group is *phenolic*.

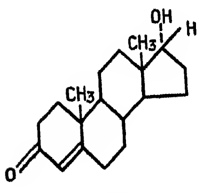
(ii) *Œstrone*: 3-hydroxy-17-keto-œstratriene.

(iii) *Œstriol* (-triol=trihydroxy): the three OH groups are in positions 3, 15, and 17.

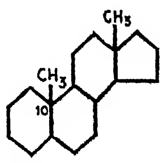
ANDROGENS (TESTOIDS) (Fig. 664).—These are based on *androstane*, which is œstrane with a second methyl, *i.e.* (CH_3) group attached to C_{10}; androstane is thus 10-methyl œstrane.

(i) *Androsterone* is 3-hydroxy-17-keto-androstane.

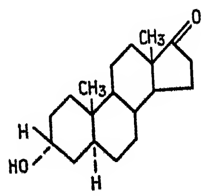
(ii) *Testosterone* is 3-keto-17-hydroxy-androstene (1 double bond joins C_4 and C_5.



Testosterone.
In testis; action on comb growth 10 × androsterone; accessory reproductive organs of mammals: very active.



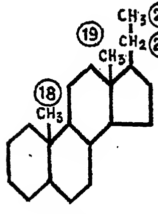
Androstane.



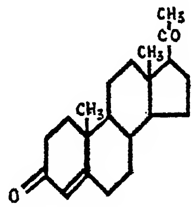
Androsterone.
In male urine. Comb growth +; accessory reproductive organs of mammals: slight action.

FIG. 664.—Structure of Androgens.

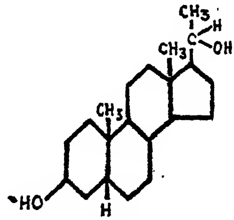
PROGESTINS (LUTEOIDS) (Fig. 665).—The basic compound is *pregnane* which resembles androstane, but has a $-CH_2-CH_3$ side chain attached to C_{17}. (The CH_3 attached to C_{10} is numbered 19; the CH_3 attached at C_{13} is numbered 18: the C atoms of the side chain attached to C_{17} are numbered 20 and 21.) *Pregnene* has one double bond joining C_4 and C_5.



Pregnane.



Progesterone.
In corpus luteum.



Pregnanediol.
Inactive. In urine.

FIG. 665.—Structure of Progestins.

(i) *Progesterone* is 3: 20-diketo-pregnene.

(ii) *Pregnanediol* is a fully reduced derivative of progesterone with 6 additional H atoms.

Œstrogens (Œstrus-producing Substances).—As œstrus does not occur in women and the higher primates, this term is a misnomer when applied in these species to œstradiol and related substances. The term is, however, sanctioned by historical reasons and common usage. The principal natural œstrogens in women are (i) *œstradiol* which is probably the hormone secreted by the ovary; (ii) *œstrone* and *œstriol*, less biologically potent substances which are derived from œstradiol and are secreted in the urine.

SOURCES.—Œstrogens can be extracted from the ovary (both the *Graafian follicle* and the *corpus luteum*); the placenta; the adrenal cortex; the testis; and normal adult *male* as well as female urine. There is a large excretion in the urine of pregnant women; surprisingly, there is also a large excretion in the urine of stallions.

PRODUCTION AND FATE OF ŒSTROGENS.—(i) Injection experiments show that about 10 per cent. of injected œstrogen is excreted in the urine; this fact is made use of to calculate the natural rate of œstrogen formation.

(ii) An adult woman excretes in the urine about 1 mg. of œstrogen per month, corresponding to the formation of 10 mg. per cycle. Œstrogen is excreted in the urine throughout the cycle, reaching its peak at about the time of *ovulation*; this is believed to be derived from the cells of the *Graafian follicle* whose secretory activity increases as it grows and matures. Œstrogen is also secreted during the second (or corpus luteum) phase of the cycle; there may be a secondary rise in œstrogen output at this time in the urine. It seems, therefore, that the corpus luteum is a *dual gland* secreting both œstrogen and progesterone. As mentioned above, œstrogen can be extracted from the corpus luteum as well as from the follicles. At the menopause, as the follicles progressively disappear, the urinary excretion of œstrogen declines and finally ceases. No œstrogen is excreted by children before puberty.

(iii) Œstrogen is excreted in increasing amounts during the last six months of *pregnancy*; when its excretion is at its peak during the last three months-

the daily œstrogen output is 15–45 mg. per day (mostly as œstriol, to a small extent as œstrone). The œstrogen of pregnancy is formed by the placenta:

(a) the amount excreted is related to the size of the placenta (Fig. 669);

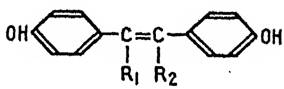
(b) excretion falls rapidly to normal levels after expulsion of the placenta.

(iv) The significance of the œstrogen content of the adrenal cortex and testis is unknown. (The adrenal cortex also contains androgen (cf. p. 968).

(v) Œstradiol is metabolized largely in the liver; it is partly oxidized to inert compounds, and partly converted to weaker derivatives like œstrone and œstriol, which are then conjugated with glucuronic and sulphuric acids and finally excreted in the urine. In *hepatic insufficiency* signs of œstrogen overactivity may appear.

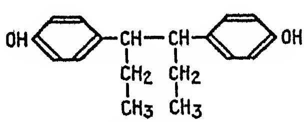
Artificial Œstrogens.[1]—Some highly potent œstrogens have been synthesized which differ in their chemical structure from the natural œstrogens. Since the natural substance œstrone has now been synthesized it is more appropriate to call these compounds "artificial" rather than "synthetic" œstrogens.

Dihydroxystilbene Compounds.—These compounds have the general formula:

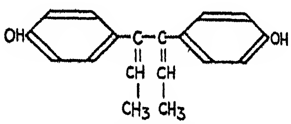


The best-known compound of this group is that in which both R_1 and R_2 are C_2H_5 groups *diethyl) stilbœstrol,* or more simply *stilbœstrol).*

Other closely related compounds are hexœstrol, benzœstrol, and dienœstrol.

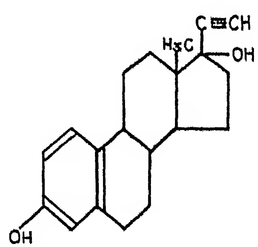


Hexœstrol.



Dienœstrol

Ethinyl Œstradiol.—This synthetic œstrogen differs from those mentioned above in that it is merely a modification of the natural hormone, œstradiol.



Ethinyl œstradiol.

Unlike the natural œstrogens the artificial substances are highly active when given by mouth. Œstradiol, as stated above, is rapidly metabolized by

[1] Dodds, *Quart. J. Pharmacy Pharmacol.*, 1949, *1*, 137.

the liver; when administered orally and absorbed via the portal vein, it is largely inactivated during its passage through this organ. Stilbœstrol is much more resistant to such changes, though a small fraction appears in the urine as the monoglucuronide. Stilbœstrol also has the advantage of being cheap to prepare.

The relative œstrogenic potencies of these different substances vary greatly according to the test used. The most reliable comparison can be made in ovariectomized women by determining the total dose which must be given for 2 weeks to produce "œstrogen withdrawal bleeding" (p. 1072). In this way it has been shown that stilbœstrol is 4 times as potent as dienœstrol, and 18 times as potent as hexœstrol; stilbœstrol, however, is more liable than the other compounds to produce nausea in therapeutic doses. Ethinyl œstradiol is the most potent œstrogen known, and is about 50 times as powerful as stilbœstrol.[1]

Actions of Œstrogens.—(1) Relation to Œstrus.—In the *lower mammals* the outstanding effect of œstrogen is to produce the uterine, vaginal, and other somatic and psychic effects characteristic of œstrus (p. 1074).

(2) Action on Uterus and Adnexa in Women.[2]—In *ovariectomized* women undergoing œstrogen treatment the uterine muscle fibres increase in number and size and their spontaneous motility is enhanced; the atrophied mucosa develops; there is also some increase in the size and vascularity of the *cervix*. The secretory activity of the cells lining the *Fallopian tubes* is stimulated and *the motility* of its muscle coat and of the cilia is increased. Stratification of the *vaginal epithelium* is promoted.

(3) Relation to Menstrual Cycle.—In ovariectomized women the injection of large doses of œstrogen produces changes which are identical with the *proliferative* phase of the menstrual cycle. When œstrogen injections are *stopped, uterine bleeding occurs.*

In some monkeys reddening and swelling of the perivulval (sexual) skin is produced; fluid is also retained in the *skin* generally.

(4) Breast.—Œstrogen produces various degrees of proliferation of the *ducts* and occasionally of the alveoli of the breasts (p. 1092); it also induces growth of the epithelium of the *nipple* in both sexes. Cancer of the breast may be produced in susceptible mice (see footnote 1).

(5) Action on Anterior Pituitary.—Œstrogen inhibits the secretion of the anterior pituitary hormones, including the gonadotrophins. In large doses œstrogen may thus *indirectly* cause atrophy of the ovaries and consequently arrest of menstruation, atrophy of the testes and sterility in both sexes. Œstrogen has no *direct* action on the ovary or testis. The normal regulation of gonadotrophic secretion by œstrogen is considered on p. 1085.

(6) Rôle in Pregnancy and Parturition.—See pp. 1087, 1091.

(7) Therapeutic Uses.—See p. 1082.

Functions of Corpus Luteum. Progestins. Progesterone.—The

[1] Many substances have been extracted from tar or prepared synthetically which are carcinogenic, *i.e.* when painted repeatedly on the skin of the rat they produce cancer; though these substances differ markedly chemically from the natural œstrogens they are also capable in massive doses of inducing œstrus. Conversely, large doses of œstrogen may produce cancer of the breast in susceptible strains of mice. There is no reason to suppose, however, that the doses of œstrogen used clinically may be followed by this disastrous result in women.

[2] Similar results are obtained in the primates in which a menstrual cycle occurs.

corpus luteum, through its active principle *progesterone* is responsible for:

(i) The *premenstrual growth phase* in the non-pregnant human uterus.
(ii) Certain *breast* changes.
(iii) The development of the *placenta* (decidua) during pregnancy and the embedding of the fertilized ovum in the uterus.
(iv) Certain other changes *during pregnancy*.

1. Relation of Progesterone to Menstrual Cycle.—In the normal cycle the secretion of progestrone begins two days after the occurrence of ovulation and stops about two days before the onset of menstruation. During these 10–12 days, 50–100 mg. of progesterone are secreted of which about one-quarter is excreted in the urine as Na pregnanediol glucuronidate. The rôle of progesterone on the menstrual cycle is considered on p. 1073.

2. Functions During Pregnancy.—(1) Placenta Formation and Embedding of Ovum.—The relation of the corpus luteum to these functions is illustrated by the following experiments:

(i) As stated on p. 1084, ovulation in the rabbit only occurs after coitus; following rupture of the follicles, corpora lutea develop. If copulation is carried out by sterilized bucks, corpora lutea are formed, but, of course, *no pregnancy* ensues. During the period of persistence of the corpus luteum the uterus shows growth changes in its mucosa, similar to those of early pregnancy. If a mechanical stimulus is applied to the uterine mucosa (simulating the presence of a fertilized ovum) very marked local growth occurs, leading to the development of a decidua, *i.e.* placenta formation has been artificially induced. The corpus luteum thus forms an internal secretion which sets up the growth ("progestational") changes in the mucosa that are necessary for the embedding and development of the fertilized ovum.

(ii) If the corpora lutea are removed from rabbits 20 hours after mating (when the fertilized ova have been in the Fallopian tubes for about 10 hours), no progestational changes occur in the uterus, and none of the embryos live after the fourth day. The above experiment is then repeated, but after the ovariectomy, extracts of corpus luteum (or progesterone) are administered for 5 or 6 days. The *embryos persist*, and histological examination of the uterus at the end of this period shows *progestational changes* identical with those seen in normal pregnancy. If the injections are continued long enough the pregnancy may go on to term. [Œstrogen has no such effects.]

The presence of the fertilized ovum in the uterus causes the corpus luteum to persist till the end of pregnancy. The responsible agent is chorionic gonadotrophin which is formed by the placenta (it is also responsible for the Aschheim-Zondek reaction and the other diagnostic tests of pregnancy (p. 1088)).

(2) Action on Uterine Muscle.—Corpus luteum extracts (*progesterone*) in *rabbits* reduce the spontaneous activity of the uterine muscle and its responsiveness to oxytocin; the uterus can thus enlarge progressively to accommodate its expanding contents during pregnancy. If the corpus luteum is made to persist too long by injections of anterior pituitary extracts the onset of parturition is delayed. On the basis of these experiments it has been suggested that degeneration of the corpus luteum at the end of pregnancy may be a necessary condition for the onset of labour. There may, however, be important species differences, as in *women* injection of

progesterone may induce slow uterine contractions of high amplitude. The progesterone secreted during the latter part of pregnancy in women would then *not* induce uterine quiescence, nor would the cessation of progesterone secretion help to promote parturition.

In women the *placenta* also secretes progesterone and is thus concerned with its own growth (p. 1088). If ovariectomy is carried out in women early in pregnancy the fœtus may go to term in spite of the absence of the secretion of the corpus luteum.

(3) Action on Breasts.—The corpus luteum is one of the factors concerned with the growth of the alveolar tissue of the breasts during pregnancy (p. 1092).

(4) Inhibition of Ovulation.—During pregnancy, ovulation (and menstruation) do not take place; this is probably related to the persistent secretion of progesterone. Experimentally, injection of progesterone prevents ovulation and delays the onset of menstrual bleeding; it is not known how it produces the former effect.

"Relaxin" and Pelvic Ligaments.—In the latter part of pregnancy the pelvic ligaments relax "in preparation" for the stretching which accompanies labour. This response is attributed to a substance called "relaxin," which, though not yet chemically identified, is distinct from œstradiol and progesterone. It is formed mainly by the placenta, and to a much smaller extent by the uterus, from which it may be liberated into the blood stream by the actions of progesterone or œstradiol.

3. **Chemistry of Corpus Luteum Group** (Fig. 665).—This subject has been considered on p. 1077. The active principle which has been extracted from the corpus luteum is *progesterone*. It has been synthesized from stigmasterol (a sterol found in soya bean) which is closely related to cholesterol. An inactive derivative—*pregnanediol*—is found (as the glucuronide) in the urine of pregnancy, rising in amount from the eighth week to about eightfold at parturition. It also appears in the urine during the second half of the normal menstrual cycle and disappears again just before the onset of menstruation. It is an index of the secretion of progesterone about 25% of which is converted to pregnanediol.

Progesterone-like substances have been synthesized, the most important being ethisterone (pregneninolone); these compounds have the advantage of being active when administered by mouth (unlike progesterone). The formula of ethisterone is shown in Fig. 666 and should be compared with that of progesterone (p. 1077).

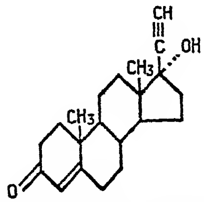


Fig. 666.—Structure of Ethisterone.

Progesterone is related chemically to certain of the adrenal corticoids (Fig. 610) and may preserve life in adrenalectomized animals.

Therapeutic Uses of Sex Hormones.—1. Œstrogens.—The substance of choice clinically is *stilbœstrol* (p. 1078) which is cheap and can be given by mouth. The therapeutic uses of œstrogens are summarized below.

(1) *Menopausal Symptoms.*—Œstrogen controls satisfactorily the hot flushings which are a common and troublesome symptom. If suitably administered together with progesterone a normal menstrual cycle can be produced in ovariectomized women; as amenorrhœa itself is harmless, little advantage is gained by the patient from this interesting physiological experiment.

(2) *Delayed Puberty.*—The age of onset of menstruation and sexual maturity in girls in Britain is 10–15 years; but 25% of girls do not menstruate till they are 16 years of age or over. If the sexual organs are still immature at the age of 18 it may be advisable to administer œstrogen to promote the growth of the persistently infantile uterus and of the breasts.

(3) *Vulva and Vagina.*—Œstrogen increases the thickness and degree of stratification of the vulval and vaginal epithelia; it has proved of value in allaying *vulval itching* and especially that associated with *kraurosis vulvæ* and *senile vaginitis.* Combined with sulphonamide, œstrogen may be of use in the treatment of gonorrhœal *vulvo-vaginitis.*

(4) *Nose.*—Œstrogen modifies the mucous membrane of the nose and has been useful in the treatment of *atrophic rhinitis.*

(5) *Breasts.*—If given before the onset of lactation œstrogen may inhibit the secretion of milk by the breasts; once lactation is well established it is less effective. [*Androgens* act similarly.]

(6) *Carcinoma of Prostate.*—The use of œstrogens in this condition is considered on p. 1116. Œstrogens do not benefit cases of *simple hypertrophy* of the prostate (cf. p. 1117).

(7) *Abortion.*—Stilbœstrol (but not natural œstrogen) given early in pregnancy interferes with the implantation of the ovum and may produce abortion.

(8) *Menstrual Disturbances.*—Œstrogens are commonly used empirically in a variety of ill-understood menstrual disorders in the hope that some benefit may result. Large doses may have a dramatic hæmostatic effect in patients with prolonged and heavy uterine bleeding.

2. Progesterone Group.—Members of the progesterone group are sometimes useful in cases of repeated abortion. As pointed out on p. 1091, progesterone in *women* may induce slow contractions of high amplitude. Its claimed beneficial effects in preventing abortion may therefore be due to its capacity to stimulate the *development of the decidua* and the *formation of the placenta* in patients in whom progesterone secretion is deficient.

3. Androgens.—The action of the androgens in the female may be conveniently referred to here. They may be: (i) androgenic, *i.e.* masculinizing, producing, for example, growth of the clitoris; (ii) gynæcogenic, *i.e.* feminizing, promoting growth of the infantile uterus and causing œstrous changes in the uterus; (iii) in *very large doses* they may depress the pituitary gland and thus indirectly inhibit ovulation and menstruation and influence the breasts. Clinically the androgens have proved beneficial in cases of excessive uterine bleeding (menorrhagia); large doses of testosterone may have temporary beneficial effects in *carcinoma of the breast,* acting particularly on bony metastases by causing sclerosis; like the œstrogens the androgens may

inhibit lactation. Administration of these substances is not free from disadvantages, as they may cause growth of a beard and breaking of the voice (*i.e.* general phenomena of masculinization).

RELATIONSHIP OF ANTERIOR PITUITARY TO OVARY. THE GONADOTROPHINS

Anterior Pituitary and Ovary.—The anterior pituitary by means of its gonadotrophic hormones initiates the cyclic changes in the ovary at puberty and regulates these changes throughout active sexual life. It causes the ovary to secrete its hormones, *i.e.* œstrogen and progesterone. Indirectly, therefore, the pituitary controls the structure and functions of the rest of the reproductive apparatus and the secondary sexual characters in the female.

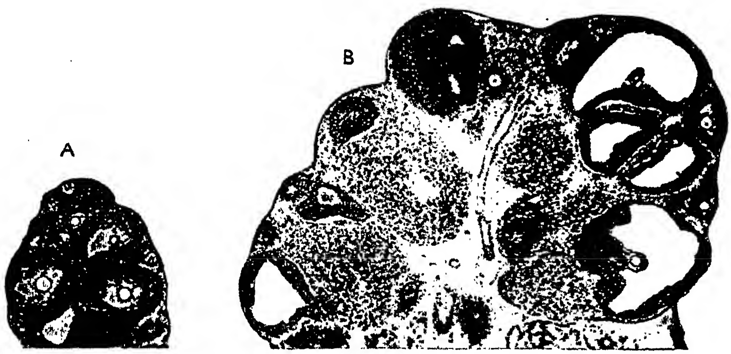


FIG. 667.—Action of Anterior Pituitary Transplants on Immature Ovary in Rat. (Smith and Engle, from Parkes, *Internal Secretions of the Ovary.*)

A and B are ovaries of litter-mate, immature rats.
A = untreated.
B = received four daily transplants of anterior pituitary material.
Note the great increase in the size of the ovary of B and the formation of many mature Graafian follicles

(1) RESULTS OF HYPOPHYSECTOMY.—Hypophysectomy in immature animals causes the ovaries to remain infantile; in adult animals cyclic ovarian activity ceases. Anterior pituitary transplants in immature animals cause rapid growth of the ovary and precocious onset of ovulation and corpus luteum formation (Fig. 667); many ova are discharged. In adult hypophysectomized animals pituitary transplants restore cyclic ovarian activity.

(2) PITUITARY GONADOTROPHINS.—The anterior pituitary acts on the ovary by means of three hormones:

(i) *Follicle-stimulating hormone* (FSH).

(ii) *Luteinizing hormone* (LH) which is chemically identical with interstitial-cell-stimulating hormone (ICSH) that acts on the interstitial cells of the testis. For this reason LH is often also referred to as ICSH.

(iii) *Prolactin* (*lactogenic hormone, lactogen, luteotrophin*).

FSH and LH (ICSH) are glycoproteins; they are the classical gonadotrophins, *i.e.* hormones acting on the gonads. The action of prolactin was first thought to be limited to the breast, hence its name. It is now established,

however, that it initiates and maintains the *secretory activity* of the corpus luteum, hence its new name, luteotrophin. It should not be confused with luteinizing hormone which is in part responsible (together with FSH) for the *formation* of the corpus luteum. Prolactin is obviously a third gonadotrophin; it is a simple protein.

(3) Other Gonadotrophins.—Two other gonadotrophins have been isolated :

(i) *Chorionic gonadotrophin*, which is present in the urine of pregnant women ; it is formed in the placenta. It has similar physiological actions in women to LH but differs from LH in structural details. It also has a luteotrophic action.

(ii) *Pregnant mare serum gonadotrophin* (*PMS*, *serum gonadotrophin, equine gonadotrophin*).—This substance has not yet been isolated in the pure state ; it resembles FSH in its action in women.

(4) Action of Gonadotrophins.—The actions of the individual gonadotrophins are best demonstrated by injecting them into hypophysectomized immature animals.

(i) FSH produces growth of the Graafian follicle. Given alone it does not induce ovulation.

(ii) If FSH is now given combined with LH, and the dosage of the two hormones is appropriately adjusted, further growth of the follicle takes place, ending in the occurrence of ovulation and the formation of the corpus luteum.

(iii) If prolactin is now given, the secretory activity of the corpus luteum is initiated and sustained. On stopping the injections the corpus luteum atrophies.

There must obviously be some mechanisms in the normal animal to regulate the timing of the release of the individual gonadotrophins and the dose of hormone so released.

Regulation of Discharge of Gonadotrophins.—The secretion of these hormones is regulated by :

(i) the hypothalamus ;
(ii) the concentration of ovarian hormones (mainly œstrogen) in the blood.

(1) Rôle of Hypothalmus.—(i) In the rabbit electrical stimulation of the hypothalamus produces ovulation. As explained on p. 931 a chemical transmitter is released locally which is carried in the blood to the anterior pituitary causing a discharge of gonadotrophins.

(ii) In the rabbit, ovulation does not normally occur spontaneously and rhythmically, but only takes place 12–18 hours after coitus. In the absence of the anterior pituitary, however, coitus does not produce this effect. The act of coitus in the rabbit probably reflexly stimulates the hypothalamus; the further sequence of events is as in (i) *supra*.

(iii) The winter period is usually a non-breeding season in the ferret; in the female the ovaries are quiescent during this time and do not show the characteristic cyclic changes. But reproductive activity sets in even in winter if the number of hours of light to which the animals are exposed daily is prolonged by artificial illumination. The light stimulus acts on the retina, sends excitatory impulses up to the brain, and presumably ultimately stimulates the hypothalamus. The effects of light cannot be produced after hypophysectomy and therefore involve the secretion of gonadotrophins.

(iv) A similar reflex comes into play in the female pigeon; it does not ovulate spontaneously but does so usually in the presence of a male bird. It will also ovulate, however, in the presence of another female; in fact the bird will ovulate if left alone in a cage so long as it is provided with a mirror in which it can see its own reflection. Here, too, visual stimuli lead (presumably via the hypothalamus) to the discharge of gonadotrophins. [The effects of self-examination in women are more generalized in character.]

(2) Rôle of Blood Hormone Level.—(i) When œstrogens are injected

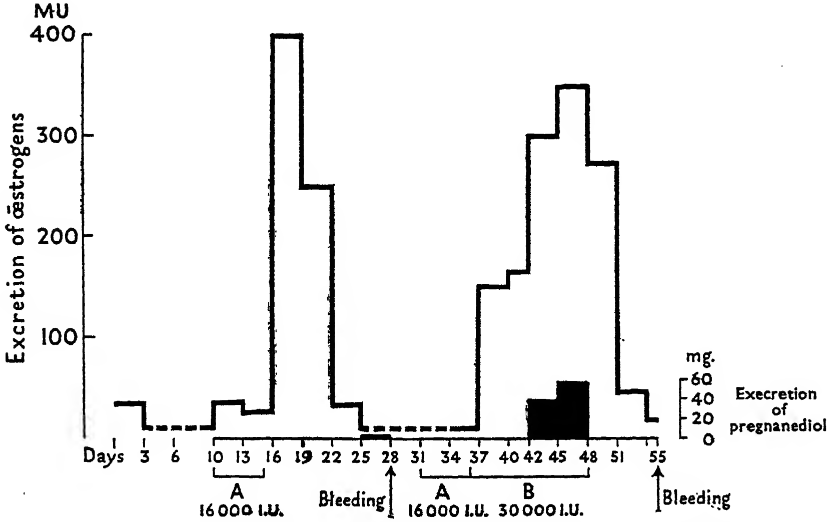


Fig. 668.—Effect of Serum Gonadotrophin (=FSH) and Chorionic Gonadotrophin (=LH plus Luteotrophin) on Urinary Excretion of Œstrogens and Pregnanediol in Case of Primary Amenorrhœa.

Clear columns: excretion of œstrogens (in mouse units).
Black columns: excretion of pregnanediol (derived from progesterone) in mg.
A: Inject serum gonadotrophin (FSH).
B: Inject chorionic gonadotrophin (LH plus luteotrophin).
After A, bleeding occurs from an endometrium in the proliferative phase.
After B, bleeding occurs from an endometrium in the progestational phase.
(Redrawn from Rydberg *et al.*, *J. Amer. med. Assoc.*, 1943, *121*, 1121.)

for long periods, they produce marked structural and functional changes in the anterior pituitary. The gland becomes considerably enlarged or develops an appearance resembling an innocent pituitary tumour; it may show intense congestion and hæmorrhages or other destructive changes. The functions of the gland are often gravely deranged; thus the secretion of gonadotrophic hormones may be inhibited in both sexes leading to atrophic changes in the ovaries or testes. [Secretion of growth hormone may also cease, leading to dwarfism. Œstrogens can thus profoundly modify many anterior pituitary activities.]

(ii) At the menopause, when the ovary ceases to secrete œstrogen and progesterone, there is increased excretion in the urine of FSH.

(iii) Injection of œstrogen in suitable doses may inhibit the secretion of FSH and increase the secretion of luteinizing hormone and of prolactin.

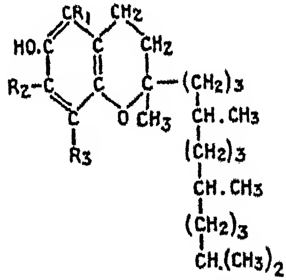
These scattered observations suggest that there is probably a complex *iprocal relationship* between the ovary and the pituitary. It is supposed .t at the onset of menstruation, when the blood level of œstrogen and pro-.terone falls to a minimum, FSH is secreted causing the initial follicular owth. The resulting increased œstrogen level decreases the rate of secretion FSH and stimulates the secretion of LH and presumably of prolactin, using ovulation, luteinization, and the secretion of progesterone. There is :le information about the action of progesterone on the pituitary.

Clinical Use of Gonadotrophins.—The gonadotrophins used are *chorionic gonadotrophin* (derived from human pregnancy urine), which ts like LH plus luteotrophin; (ii) *serum gonadotrophin*, derived from egnant mare serum which acts mainly like FSH. In cases of *primary ienorrhœa* in which the ovaries are thought to be in the infantile state, id more commonly in *secondary amenorrhœa*, injected gonadotrophins may imulate ovarian activity and secondarily uterine development and thus 'oduce normal menstrual cycles. In women with infantile ovaries, serum onadotrophin (=FSH) given alone stimulates ovulation and œstrogen libera-on; the uterine mucosa passes into the proliferative phase; there is a great crease in œstrogen excretion in the urine. When this treatment is followed y injection of chorionic gonadotrophin (=LH+luteotrophin), luteinization the ovary takes place; sometimes the ovary becomes greatly enlarged id tender owing to the formation of a "lutein tumour"; progesterone is creted and the uterine mucosa passes into the progestational phase; regnanediol excretion in the urine rises markedly. Some 10 days after the st injection menstruation sets in (Fig. 668). Some previously sterile women ave been helped to become pregnant by such treatment.

RELATION OF DIET TO SEXUAL FUNCTION. VITAMIN-*E*[1]

Experimental work in the *rat* has shown that in this animal two of the itamins (-*E* and -*A*) are related closely to normal functioning of the eproductive apparatus. It is not known whether the same is the case in the igher mammals.

(1) **Vitamin-*E*** is found in highest concentration in the embryo f *wheat*, in most commercial oils (*e.g.* olive oil), in seeds and green eaves; there is little in milk fat and none in cod-liver oil. It is present, but not in high concentrations, in animal tissues, chiefly in the at and muscles. Chemically it is a *tocopherol* with the constitution ndicated below.



[1] Symposium on Vitamin-*E*, *Ann. N.Y. Acad. Sci.*, 1949, *52*, 63–428.

R_1, R_2, and R_3 are alkyl groups. Several forms of tocopherol have been isolated. In α-tocopherol, $R_1=R_2=R_3=CH_3$.

Effects of Absence of Vitamin-E.—In the *male rat* atrophy of the seminiferous tubules eventually occurs. In the *female rat*, œstrus, ovulation, coitus, and implantation of the ovum in the uterus occur normally. The embryo develops normally until the eighth day; further growth proceeds slowly and before the twentieth day death of the fœtus occurs. The maternal part of the placenta may go on growing and the mother increases in weight, suggesting that the vitamin is essential only for fœtal development. The embryo shows maldevelopment of blood vessels, blood cells, yolk sac, and allantois. If the vitamin is given on the fifth day of pregnancy the fœtus can still be saved. If animals are fed on diets rich in -*E* and then placed on diets from which the vitamin is absent, they retain fertility for three or four months. Excess of vitamin-*E* does not increase fertility beyond the normal.

Vitamin-*E* concentrates have been administered during pregnancy to women suffering from habitual abortion; though claims have been made that such treatment was of benefit, many observers are unimpressed with the results obtained.

Vitamin-*E* deficiency in the rat is said to lead to the development of *muscular dystrophy* which can be cured by administration of α-tocopherol.[1] The vitamin has been used with occasional success in the treatment of clinical myopathies.[2]

(2) *Vitamin-A.*—When the diet is deficient in vitamin-*A*, sexual maturity is delayed and ovulation is infrequent; there is extensive keratinization of the vagina and continuous shedding of horny cells which obscures the normal changes which occur during the œstrous cycle (cf. p. 1074).

PHYSIOLOGY OF PREGNANCY. PARTURITION.

Physiological Changes in Pregnancy.[3]—The principal changes taking place as a result of pregnancy are considered in various parts of this book. The more important matters dealt with are the following:

(i) Development of the *placenta* (p. 1080). It must be emphasized that the placenta is also an organ of *internal secretion*, forming in women *chorionic gonadotrophin œstrogen*, and *progesterone* (see below and Figs. 669, 670). The rôle of the placenta in *fœtal respiration* is dealt with on pp. 1097 *et seq.*

(ii) Increase in the size of pregnant uterus, relaxation of the pelvic ligaments (p. 1081), and enlargement of the birth canal.

(iii) Formation, persistence, and degeneration of the corpora lutea (p. 1067), and cessation of ovulation and menstruation (p. 1081).

(iv) Development of the *breasts* (p. 1092) and the onset of *lactation* (p. 1093).

Excretion of Hormones in the Urine during Pregnancy.—(1) Œstrogen (Fig. 669).—There is a progressive increase in the urinary œstrogen (*secreted by the placenta*) which rises to its peak just before the onset of parturition. It is mainly in the form of œstrone combined with

[1] Einarson and Ringsted, *The Effects of Chronic Vitamin-E deficiency on the Nervous System and Skeletal Musculature in Rats*, Copenhagen 1938.

[2] Rabinovitch *et al.*, *J. Neurol. Neurosurg. Psychiat.*, 1951, *14*, 95.

[3] Newton, *Physiol. Rev.*, 1938, *18*, 419; *Recent Advances in Physiology*, 7th edn., London, 1949, p. 86.

lucuronic acid; after delivery, the total excretion rapidly declines. During regnancy the blood œstrogen concentration also rises and follows the same eneral course as the excretion in the urine (see p. 1077).

(2) PREGNANEDIOL.—The excretion of progesterone in the urine in the orm of pregnanediol glucuronide rises. At the eighth week of pregnancy it s about 10 mg. per 24 hours; at the end of pregnancy it averages 80 mg. Fig. 670). The progesterone is derived mainly from the placenta and partly erhaps from the corpus luteum. Pregnanediol excretion persists after ovariectomy.

(3) CHORIONIC GONADOTROPHIN.—During pregnancy the placenta in *women* forms chorionic gonadrotrophin (p. 1084). It appears in the urine as early as eight days after the first missed period, when the pregnancy has lasted

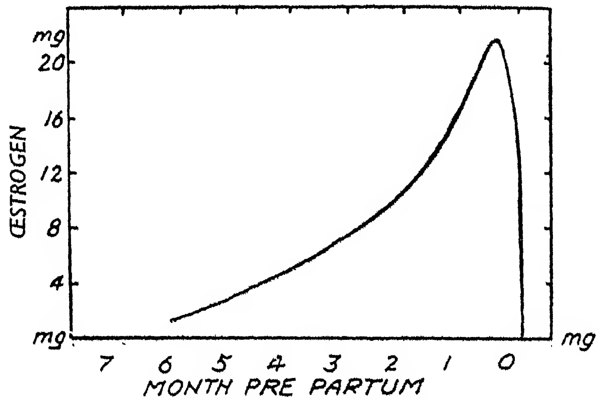


FIG. 669.—Excretion of Œstrogen in Urine in Pregnant Women.

Ordinate: mg. per 24 hours.
(Data of Marrian *et al.*, from Robson, *Recent Advances in Sex and Reproductive Physiology*, Churchill.)

for less than a month. It reaches its peak at 50–60 days, falls to a very much lower level at 80 days, then remains steady and disappears a few days before parturition (Fig. 670). It persists during the puerperium if placental tissue is retained in the uterus; it disappears from the urine if the fœtus dies. The presence of chorionic gonadotrophin in the urine is, apart from certain exceptional conditions, convincing evidence of pregnancy; it is the basis of the various tests of pregnancy (p. 1089).

Chorionic gonadotrophin appears in the urine in large amounts in cases of *hydatidiform mole* and especially of the malignant tumour of the placenta known as *chorion-epithelioma*. Persistence of urinary chorionic gonadotrophin after excision of the tumour is evidence of incomplete removal, and its return in the urine indicates recurrence of the growth. Cases of *teratoma of the testis* in which the tumour contains chorionic tissue also show a high urinary excretion of chorionic gonadotrophin.

Large amounts of this gonadotrophin can be extracted from normal chorionic tissue, or that grown in tissue culture or found in placental tumours

The *physiological action* of chorionic gonadotrophin depends on the species on which it is tested.

(i) Injected into women it acts like pituitary LH plus luteotrophin ; the placenta probably secretes this hormone and thus helps to maintain the activity of the corpus luteum during the first six months of pregnancy. In cases of chorion-epithelioma the ovaries often contain multiple large luteinized follicles with imprisoned ova.

(ii) In rodents, on the other hand, injection of chorionic gonadotrophin stimulates *all* the ovarian processes (follicular growth, ovulation, and corpus

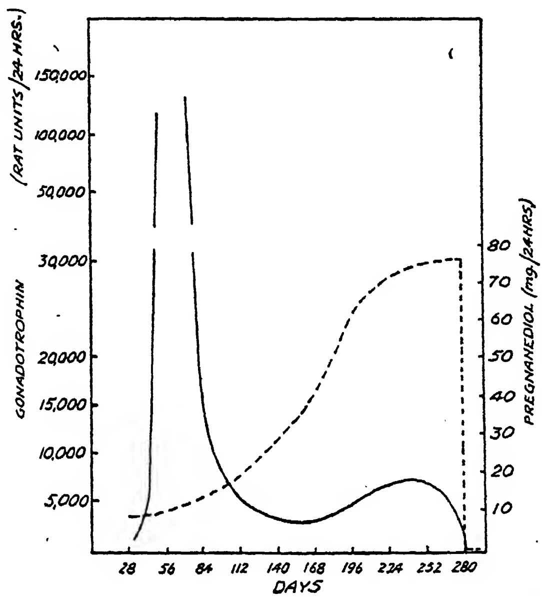


FIG. 670.—Urinary Excretion of Chorionic Gonadotrophin and Pregnanediol in Pregnant Women.

Continuous line : gonadotrophin. Interrupted line : pregnanediol.
(Robson, *Recent Advances in Sex and Reproductive Physiology*, Churchill.)

luteum formation) in the *intact animal* ; in the hypophysectomized rodent it produces only luteinization.

Pregnancy Diagnosis Tests.—The high concentrations of chorionic gonadotrophin which appear in the urine in the early stages of pregnancy form the basis of the most reliable tests for the diagnosis of pregnancy. The accuracy of these tests is of the order of 99%.

(1) ASCHHEIM-ZONDEK TEST.—The chorionic gonadotrophin of pregnant women produces a *precocious full ovarian cycle in immature mice.* The technique is as follows : small quantities of urine are injected several times daily for 3 days into immature mice (3—4 weeks old). The animals are killed and examined on the fifth day. The essential criteria for positive result (*i.e.* pregnancy present) are the presence of blood-filled follicles (corpora hæmorrhagica) or corpora lutea.

(2) Friedman Test.—Human pregnancy urine produces ovulation 18 hours after injection into virgin *rabbits* (in the absence of the normal stimulus of copulation) (cf. p. 1080).

(3) Tests in Toads and Frogs.—The first satisfactory test in these species was that described for the female of *Xenopus lævis* (the South African clawed toad). Urine is injected into the dorsal lymph sac; the presence of chorionic gonadotrophin is indicated by the shedding of numerous ova within 24 hours.

Subsequently *male* toads (and even frogs) of many varieties have been used, since chorionic gonadotrophin also acts on the seminiferous tubules and causes discharge of *spermatozoa* into the urine within 3 hours of injection.

(4) Rat Ovarian Hyperæmia Test.—Chorionic gonadotrophin produces conspicuous hyperæmia of the ovary of immature rats within 6 hours of injection.

Tests (3) and (4) certainly give the speediest answers to the question at issue.

Urinary Excretion of Hormones.—A short summary follows of the present state of knowledge of the presence of hormones in the urine.

(1) Sex Hormones in Urine.—(i) Derivatives of *œstradiol*: œstrone and œstriol (pp. 1077, 1087).

(ii) Derivatives of *testosterone*: androsterone and isodehydroandrosterone (cf. p. 1110). Urinary androgens are also derived from the adrenal cortex (cf. p. 958).

(iii) Derivative of *progesterone*: pregnanediol (pp. 1080, 1088).

The compounds in the urine have a much lower physiological potency than the related substance formed by the gonad from which they are presumed to be derived; the urinary constituents are excreted in the conjugated form either as glucuronides or as sulphates.

Adult males and females excrete both œstrogen and androgen in the urine; none appears in the urine of children till the age of six, after which the concentration rises till puberty.

After castration, both men and women may continue to excrete both androgen and œstrogen in the urine, probably derived from the adrenal cortex (p. 958).

(iv) *Gonadotrophins.*—See p. 1088.

For urinary hormones in *pregnancy* see p. 1087; during the *menstrual cycle* see pp. 1077, 1080.

(2) Antidiuretic Hormone.—This hormone appears in the urine in the following circumstances: (i) in states of dehydration from water deprivation (p. 66); (ii) after various experimental conditions which increase the activity of the hypothalamo-hypophyseal system, *e.g.* emotional states, injection of nicotine, acetylcholine, or hypertonic saline (p. 55); (iii) in some cases of ascites (p. 822).

The Kangaroo-rat ("desert rat") that lives in the torrid arid areas of Arizona reduces its urinary water loss to an extraordinary degree by intense secretion of ADH. The urea concentration in the urine may rise to 3·6 M (=over 20%; maximal urea concentration in man about 4%) and the electrolyte concentration to 1·2 N (equivalent to 13% of NaCl). The ADH content of the urine may be 50 milliunits per c.c. (the maximum in laboratory rats deprived of water for 72 hours is 6 m.u./c.c.).[1]

[1] Ames and Van Dyke, *Proc. Soc. exp. Biol. Med.*, 1950, 75, 417.

(3) ADRENAL CORTICOIDS (so-called 11-oxysteroids).—See p. 954.

(4) NEUTRAL 17-KETOSTEROIDS.—See p. 952.

Control of Parturition.[1]—The neural division of the pituitary is probably one of the many factors concerned in the onset of parturition. The following observations are relevant to a general consideration of the question.

(i) *Progesterone* in experimental animals delays the onset of parturition. If the corpora lutea are made to persist longer than usual by injection of suitable anterior pituitary extracts the period of gestation is correspondingly prolonged. But as has been emphasized, in women injected progesterone *sets up* uterine contractions (p. 1082).

(ii) *Œstrogen* injected experimentally even in huge doses has no influence on the course of pregnancy.

(iii) *Rôle of Nervous System.*—Parturition can occur after division of the spinal cord in the mid-thoracic region or after section of the sympathetic nerve supply to the uterus.

(iv) *Oxytocin.*—Pregnant animals with experimental diabetes insipidus due to hypothalamic lesions have difficulty in delivering their young (*dystocia*). They may be unable to expel the uterine contents at all, or only partially, or may die of greatly prolonged labour. The atrophic neural division in this condition is almost completely devoid of oxytocic content. The secretion of oxytocin is probably under nervous control; in these experimental animals no oxytocin is being secreted, and the failure of parturition may be due to lack of essential oxytocic hormone. It is not yet known whether the amount of oxytocin secreted is increased at the time of parturition. In some species the sensitivity of the uterine musculature to the stimulating action of oxytocin increases towards the end of pregnancy.

The cause of parturition remains a mystery.

MAMMARY GLANDS. SECRETION OF MILK[2]

Structure of Mammary Gland.—The mammary gland consists of a series of *ducts*, which branch to give rise to *terminal tubules*; these in turn lead to the *alveoli*. Covering the external surface of the epithelium of the alveoli and ducts are numerous elongated, branching, longitudinally striated cells which constitute what has been called *myoepithelium*.[3] The presumed contractile function of these cells is discussed on p. 1094.

The breast arises as an invagination from the surface epithelium which dips down into the underlying connective tissue as solid columns of cells; these gradually become hollowed out to become ducts. At *birth* the breast is rudimentary,[4] and consists essentially of the tiny nipple from which radiate

[1] Reynolds, *Physiology of Uterus*, London, 2nd edn., 1950.

[2] Loeb in Cowdry, *Special Cytology*, vol. ii., New York, 1928. Turner in Allen, *Sex and Internal Secretions*, 2nd edn., Baltimore, 1939. American Medical Association, *Glandular Physiology and Therapy*, Chicago, 1942. Folly and Malpress, in *The Hormones*, 1948, *1*, 695, 745. Symposium on "Lactation: Function and Product," *Brit. med. Bull.*, 1947, *5*, 123.

[3] Richardson, *Proc. roy. Soc. B.*, 1949, *136*, 30.

[4] Secretion of a fluid which resembles colostrum (cf. 1095) may occur in the newborn; it is attributed to slight growth and activity of the glands of the fœtus just preceding birth due to maternal hormonal influences.

a few ducts. Little further development occurs until the time of puberty. The changes which occur at *puberty* in the female vary considerably with the species studied ; in many, including the human subject, there is considerable growth and branching of the duct system ; in others there may also be formation of glandular tissue. With the recurrence of each sexual (menstrual or œstrous) cycle the gland undergoes further proliferative changes ; though this is followed by some degree of regression, on the whole, progressive enlargement takes place, which is due in part to increased deposition of fat. Between each *menstrual period* (in women) there is hyperæmia of the breasts, increase in the interalveolar stroma, and possibly new formation of alveoli ; these changes are transient.

During *pregnancy* the breasts enlarge greatly and become markedly changed in structure. During the *first* half of pregnancy, there is further duct development, but this is now accompanied by the appearance of many alveoli which form lobules. No milk is secreted by the gland cells at this stage. During the *second* half of pregnancy the epithelial cells swell and there is gradual initiation of secretory activity with slow accumulation of milk in the alveolar lumina. The further enlargement of the breast which takes place at this stage is not due to an increase in the mass of glandular tissue but to distension of the organ with its secretion. Massage of the breast may squeeze out some of this milk.

Control of Breast Development.—This is due to the complex action of a number of hormones ; *œstrogen* and *progesterone* are the primary agents responsible for mammary growth, but they appear to work best with the help of the *anterior pituitary* and *thyroid* glands. The details of the controlling mechanism varies a good deal with the species.

(1) *Action of Œstrogen.*—The injection of œstrogen into normal or castrated animals, male or female, causes thickening of the *nipple* and marked growth and branching of the *ducts* (Fig. 671, 2). These results probably account satisfactorily for the duct changes which normally occur at puberty. In most species œstrogen causes little or no glandular development, but in some animals, *e.g.* cows and goats, œstrogen administration can not only produce alveolar development but even secretion of milk ; these latter effects of œstrogen may be mediated via the anterior pituitary.

(2) *Action of Progesterone.*—Progesterone given alone, when the breast is undeveloped or following its growth under œstrogen treatment, produces no changes. But when given *together* with œstrogen (*i.e.* at same time), marked glandular development occurs, which may ultimately be equivalent to that attained at the end of the first half of normal pregnancy. No secretory changes, however, occur (Fig. 671, 3).

(3) *Rôle of Placenta.*—So much for the effects of injection experiments. In the pregnant animal, however, it must be borne in mind that the *placenta* forms both œstrogen and progesterone ; in fact it is probable that the placenta is the only source of œstrogen in the pregnant animal, and that none is formed by the ovary itself. Thus, if the ovaries are removed in pregnant mice and the placentæ happen to be retained, mammary development proceeds quite normally, indicating that adequate œstrogen and progesterone secretion still take place ; the same result occurs if the fœtuses as well as the ovaries are removed and the placentæ are retained. If, however, the placentæ are also aborted following ovariectomy, the breasts rapidly regress. It is clear, therefore,

that the placenta is an important organ of internal secretion in relation to breast development during pregnancy; its hormones stimulate proliferation both of ducts and glandular cells.

(4) An intact *nerve supply* is not essential for the growth of the mammary gland during pregnancy. Thus, if the breast is completely transplanted, thus severing all its nervous connections, it may grow during pregnancy, and function, although somewhat inefficiently, after parturition.

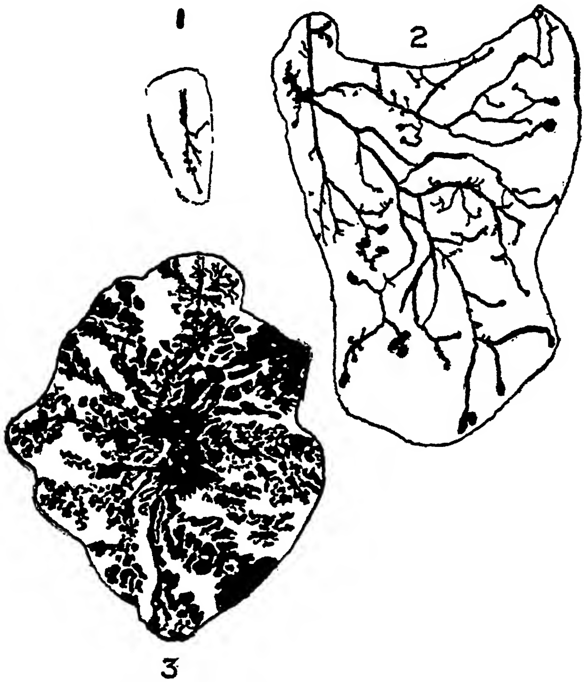


FIG. 671.—Hormonal Control of Breast Development in the Rat. (After Turner, in *Sex and Internal Secretions*, and Newton, *Recent Advances in Physiology*, Churchill.)

1. Nipple of castrated male rat.
2. Duct development produced by œstrogen.
3. Alveolar (glandular) development produced by combined treatment with œstrogen and progesterone.

Lactation.[1]—Though some secretion is present in the breasts during the latter part of pregnancy, a *free flow* of milk is established only some days after delivery of the child. It has been suggested that secretion is inhibited during pregnancy by some influence coming from the placenta or from the distended uterus; these factors would be removed at parturition. Alternatively, as explained below, the secretion of milk requires positive stimuli which only appear after parturition.

Lactation consists of two distinct processes[2]: (i) *milk secretion, i.e.* the

[1] Revised by Cyril A. Keele. [2] Folley, *Brit. med. Bull.*, 1947, *5*, 135, 142.

nthesis of milk by the alveolar epithelium, and its passage into the lumen the gland ; (ii) *discharge of milk from the breast.*

1. **Milk Secretion.**—The actual secretion of milk by the mammary and, unlike that of most externally secreting glands, is not under direct ›rvous control ; the stimulus to the alveolar epithelium comes mainly om the anterior pituitary, through its secretion of *prolactin* and other ›rmones which initiate and maintain lactation, and also from the thyroid.

(1) PROLACTIN.[1]—This pituitary gonadotrophin not only promotes pro- ›sterone secretion (p. 1083), but is also one of the important lactational ›rmones. It can only act on a breast that has been caused to grow under strogen-progesterone stimulation. It acts directly on the gland alveoli. hus if it is injected into a duct of a suitably prepared breast, lactation is ›oduced only in the corresponding alveoli and not in adjacent ones which ad into an independent duct. Prolactin causes the epithelial cells lining the veoli to increase in size and number, and many are cast off in the first milk ›creted.

There is evidence that prolactin secretion is stimulated by *suckling,* robably via a nervous pathway to the hypothalamus, and thence by means ' the neuro-humoral mechanism described on p. 931.

(2) OTHER ANTERIOR PITUITARY HORMONES.—It is probable that rolactin is optimally effective only when acting in conjunction with *lrenocorticotrophic hormone,* which in turn stimulates the adrenal cortex to ›crete its corticoids ; this fact may account for the failure of prolactin alone › increase a poor milk flow in women. *Growth hormone* may be concerned in ıe formation of the constituents of the milk.

(3) THYROID.—Normal *thyroid* function is also necessary for the main- ›nance of lactation, and in some species (including man) administration of ıyroid substance (or thyroxine) helps to restore a declining milk flow. ›dinated casein, which contains thyroxine, has been found particularly fective in raising the milk yield of cattle.

2. **Discharge of Milk.**—The discharge of milk from the mammary and depends not only on the *suction* exerted by the infant, but also on a *ontractile mechanism* in the breast which expresses milk from the alveoli into ıe ducts.

It is well known that in the cow the amount of milk present in the cisterns ıd larger ducts at the commencement of milking, is only a small fraction of ıe total quantity which can ultimately be collected. It appears that stimula- on of the teat produces, after a brief interval, a sudden rise in milk pressure ı the udder, (" let-down ") and only after this phase has occurred can the ıll milk yield be obtained. A similar rise of pressure in the ducts, called the draught," occurs in women in response to the stimulus of suckling. Both let-down " and " draught " can also be produced by injection of *oxytocin.*

The physiological sequence of events is probably as follows : stimulation ' the teat or nipple causes nerve impulses to pass (via some unknown ıthway) to the supraoptic nucleus and thence along the hypothalamo- ypophyseal tract to the neurohypophysis causing the release of oxytocin ›to the blood stream. The oxytocin is then carried to the mammary gland here it produces contraction of the myoepithelium surrounding the alveoli, ıus expelling their contained milk into the ducts, which are meanwhile kept

[1] White, *Vitamins and Hormones,* 1949, 7, 253.

open by the contraction of their longitudinally arranged myoepithelial layer. This sudden outflow of oxytocin into the blood stream probably also causes the *uterine contractions* which are known to follow suckling during the puerperium in women. Thus suckling acts by neuro-humoral mechanisms to cause both milk secretion and milk discharge, and it is easy to understand how nervous and psychological factors, acting via the hypothalamus, can influence lactation.

If milk is allowed to accumulate in the breasts, and is not removed, the gland involutes (*i.e.* regresses). The absence of suckling deprives the anterior pituitary of the stimulus which normally causes it to secrete prolactin and ACTH.

Lactation is associated with a delay in the return of the menstrual periods and temporary sterility, presumably owing to non-secretion of the other gonadotrophins (FSH and LH); but women often become pregnant again while nursing.

Composition and Properties of Human Milk.—The fluid secreted during the first three days after parturition is called *colostrum*. It is deep yellow in colour and rich in protein and salts; it is coagulated into solid masses by heat, or even spontaneously. It contains large granular bodies, called *colostrum corpuscles*, which represent either discharged alveolar cells of the gland, or else leucocytes loaded with fat. These corpuscles are abundant in the first few days, and disappear at the end of the second week.

The milk formed during the first few weeks is called the intermediate or transition milk. Mature milk appears at the end of the first month.

The accompanying table indicates the composition of colostrum, mature human milk, and cow's milk:

	Protein. g-%	Lactose. g-%	Fat. g-%	Ash. g-%	Calcium g-%
Colostrum (human) .	8·5	3·5	2·5	0·37	
Mature human milk .	1·0–2·0	6·5–8	3·0–5·0	0·18–0·25	0·03
Cow's milk (average) .	3·5	4·75	3·5	0·75	0·14

The differences between human and cow's milk are very striking; human milk contains considerably less protein, less salts, and more sugar.

(1) The *protein* content of milk is highest in colostrum (8·5%), and falls during the first few weeks (2·25%) to reach a fairly steady level of about 1·25%; it diminishes rapidly towards the end of lactation. Two proteins are found:

(i) *Caseinogen* is precipitated by weak acids; it is converted by rennin into calcium caseinate which is insoluble in water, but is easily digested by gastric juice.

(ii) *Lactalbumin*: resembles serum albumin.

In human milk there are about two parts of lactalbumin to one part of

aseinogen. In cow's milk the proportions are very different: the caseinogen s six times in excess of the lactalbumin. Allowing for the difference in the otal protein content of the two kinds of milk, it follows that cow's milk ontains about six times as much caseinogen as human milk. The caseinogen f cow's milk in the stomach forms large solid masses which are relatively nsoluble. Its exact chemical composition, too, is different from that of uman caseinogen. When human milk is treated with rennin or dilute cetic acid, fine flocculation occurs.

(2) *Fat* of milk is in the form of minute globules which are *emulsified* y the dissolved albumin: the fats chiefly present are triolein, tristearin, nd tripalmitin. Free fatty acids are only found in minute amounts; cow's nilk has about eight times as high a fatty acid content.

(3) The carbohydrate of milk is the disaccharide *lactose*.

(4) The *ash* contains *Ca*, *K*, *Na*, *Cl* and *P*, but only traces of iron: this ery *low iron content* is noteworthy. Human milk contains only 0·03% of Ca (against 0·14% in cow's milk).

(5) The *vitamin* content of milk depends on the maternal diet. For uman milk the average values are: -*A*, 300 i.u.; thiamine, 10 i.u.; -*C*, 6 mg.; *D*, 10 i.u. per 100 c.c.; the average values for cow's milk are approximately he same, but the -*C* content is lower (2 mg-%).

Origin of Constituents of Milk.—The specific constituents of milk are laborated in the gland cells from certain raw materials supplied by the lood. (i) *Lactose* is derived from the *glucose* of the plasma. (ii) *Proteins* of nilk come from the plasma *amino-acids* and proteins. (iii) *Fat* is formed artly from neutral fat of the blood and partly from acetate.

CONDITIONS AFFECTING COMPOSITION OF MILK.—Milk is richer in younger women. It is unaffected by the return of menstruation, but is adversely nfluenced by illness or by emotional disturbances.

(1) EFFECT OF DIET.—The quantity and composition of milk bear a complicated relation to the *diet*. Fundamentally a *good milk can only be formed from a good diet*. A *superabundant* diet does not increase the total yield or richness of the milk unless the protein content is increased. If the diet is *inadequate*, it is found that early in lactation the body tissues are used to form milk, which is not reduced much in amount, and weight is lost; late in lactation, however, the yield of milk is reduced. The *vitamin* content of the milk depends on the amount of these substances in the diet. Alcoholic liquors, like stout, may serve to fatten the mother, but it is very doubtful whether they improve the quality of the milk in any way.

(2) EXCRETION OF DRUGS IN MILK.—Many drugs ingested by the mother may be excreted in the milk, and it is useful to distinguish two main groups:

(1) *Those which may have actions on the suckling infant.*—(i) *Bromides* may cause drowsiness and papulo-pustular skin eruptions.

(ii) *Morphine:* addiction to morphine has been reported in children of mothers addicted to the drug.

(iii) Certain *purgatives, e.g.* aloes, phenolphthalein and calomel.

(2) *Those which have no actions on the infant.*—In a number of cases drugs may pass from the mother into the milk in measurable concentrations, but no pharmacological actions can be observed.

(i) *Sulphonamides:* the concentration in milk may equal that in blood (*e.g.* 10 mg./100 c.c.).

(ii) *Penicillin:* the milk concentration is not more than 1/10 the blood level.

(iii) *Nicotine:* heavy smoking by the mother may produce a milk nicotine concentration of 0·5 mg/L, but this does not harm the baby.

(iv) *Barbiturates*, *e.g.* barbitone and phenobarbitone.

(v) *Ethyl alcohol* appears in milk only in traces, and does not cause intoxication of the infant.

(vi) *Iodides*, *salicylates*, *quinine*, and *atropine* are only found in traces in milk.

FŒTAL RESPIRATION[1]

Fœtal Respiration.[2]—The *fœtal circulation* is briefly described on p. 331. The placenta is both a maternal and a fœtal organ. It contains large (maternal) blood sinuses receiving arterial blood from the uterine artery and returning it to the veins. Numerous fœtal chorionic villi dip into these sinuses across which gaseous interchanges take place. The probable relationship between the fœtal and maternal circulations in the human placenta is illustrated in Fig. 672.

The following analogy may usefully be drawn between the fœtal and post-natal methods of respiratory exchange. If the placenta is regarded as the

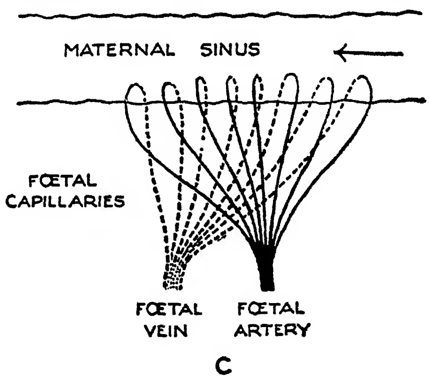


FIG. 672.—Relationship between Fœtal and Maternal Circulation in the Human Placenta. (Newton, *Recent Advances in Physiology*, Churchill.)

equivalent of the lungs, then the maternal sinuses represent the air sacs; the uterine arteries and veins, the respiratory passages; the umbilical artery and vein correspond to the pulmonary artery and vein respectively, and the chorionic villi to the pulmonary capillaries. The blood flow through the maternal part of the placenta represents the pulmonary ventilation, while the capacity of the maternal sinuses can be regarded as the equivalent of the resting content of the lungs.

The oxygen and nutritional needs of the fœtus and the amount of CO_2 and

[1] Barcroft, *Researches on Prenatal Life*, Oxford, 1946.

[2] The reader must first review the normal methods of oxygen and CO_2 transport, pp. 409, 414.

vaste products that have to be eliminated increase as pregnancy advances; n the human subject they become large during the last 3 months and especially n the last month when fœtal growth takes place most rapidly. It may)e stated that if the exchanges in the placenta are adequate to meet the 'œtal needs for *oxygen* then all the other requirements can be satisfactorily lealt with too. In the following discussion special attention will be paid ;o the problem of oxygen supply. It may be useful to recall how increased)xygen supplies are provided for the muscles during vigorous exercise, as ınalogous changes occur during pregnancy. In exercise (pp. 433 *et seq.*) the pulmonary ventilation is increased (so increasing the available O_2), the blood flow through the lungs rises (increasing the oxygen uptake by the blood) and a greatly increased blood flow is diverted to the muscles, which extract a larger fraction of the available O_2 than at rest (increased arterial-venous oxygen difference). The equivalent series of adaptations which take place in pregnancy are discussed below.

1. **Uterine Blood Flow.**—The blood flow through the maternal part of the uterus is greatly increased, perhaps finally (in the rabbit) by twentyfold; (this change corresponds functionally to an increase in the pulmonary ventilation.) In this way the oxygen supply made available for the growing fœtus is correspondingly increased. Initially, the augmented oxygen supply exceeds the needs of the still tiny fœtus, so that the blood in the uterine veins may leave almost fully saturated with oxygen, convincing testimony that the fœtus has abstracted very little. But as pregnancy proceeds, fœtal growth accelerates far more rapidly than does uterine blood flow; it is then found that a progressively greater proportion of the oxygen brought to the uterus is taken up by the fœtus and the blood which leaves in the uterine veins becomes less and less saturated with oxygen (Fig. 673). Just before parturition there is a sudden fall in the uterine blood flow which may represent a preparatory protective closing-down of the placental circulation.

2. **Fœtal Changes.**—The cardiac output of the fœtus increases throughout pregnancy in direct proportion to the increase in its body weight; the

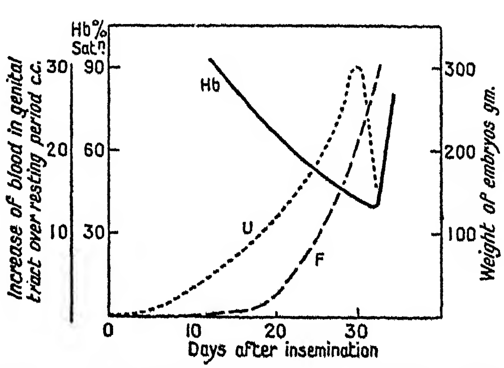


FIG. 673.—Uterine Blood Flow and Fœtal Needs in the Rabbit. (After Barcroft.)

U = blood flow through pregnant uterus; F = weight of fœtus; Hb = percentage saturation with oxygen of blood in uterine veins. Just prior to parturition (at 32 days) the uterine blood flow declines; after parturition the oxygen saturation of the blood in the uterine veins returns to normal.

blood flow (via the umbilical artery and vein) through the fœtal part of the placenta (the chorionic villi) presumably increases to a similar degree. [This corresponds to an increase in the blood flow through the lungs, which enables more oxygen to be taken up.] The maximum oxygen-carrying power (*i.e.* hæmoglobin concentration) of each unit of fœtal blood, however, remains unaffected. But, as already explained, the proportionate increase in the size of the fœtus (in the human being from the microscopic ovum to the 7-lb. baby) is much greater than that of the placenta. Though the blood supply to the placenta may increase twentyfold, the oxygen needs of the fœtus rise in an almost astronomical manner. To meet the requirements of the fœtus certain fundamental modifications occur in the oxygen-binding properties of its blood.

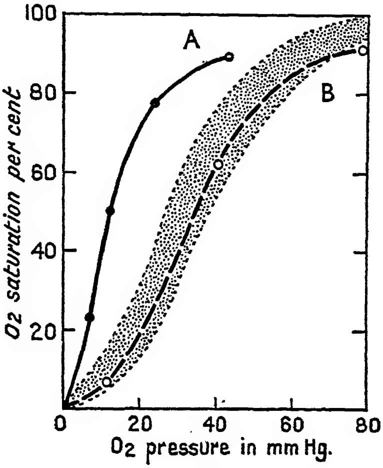


FIG. 674.—Changes in Oxyhæmoglobin Dissociation Curves in Pregnancy. (Roos and Romijn, *J. Physiol.*, 1938, *92*, 261.)

The shaded area covers the range of dissociation curves of the normal adult cow. The continuous line on the left (A) is the curve of the fœtal blood. The interrupted line on the right (B) is the curve for the maternal blood. Cow pregnant 8 months. CO_2 pressure 43–45 mm. Hg, temp. 38·5° C.

3. **Oxygen Dissociation Curve of Fœtal Hæmoglobin.**—All the hæmoglobin in the blood of the early fœtus is of the so-called *fœtal* type; it can be distinguished from *adult* hæmoglobin spectrographically, by its characteristic electrophoretic mobility, and in other ways. The adult type begins to appear in the blood at mid-pregnancy when the bone marrow begins to function as a hæmopoietic organ. It forms 6% of the total circulating hæmoglobin at the 20th week of pregnancy, 20% at birth, 50% at 2 months post-natal, and 90% at 4 months.[1] The last two values suggest that the fœtal type of hæmoglobin is not formed after birth and that the corpuscles containing it are destroyed during the first 4 months or so of post-natal life, as would be expected from the known survival time of circulating red cells (p. 186).

The outstanding functional characteristic of fœtal hæmoglobin is that its oxygen dissociation curve shows a marked *shift to the left* compared with the curve for adult maternal hæmoglobin (Fig. 674); (this difference is not the result of differences in *p*H value or CO_2 pressure). As a result, fœtal blood can take up much larger volumes of oxygen than adult blood at low O_2 pressures. Thus at 20 mm. O_2 pressure fœtal hæmoglobin is 70% saturated; at 40 mm. it is 90% saturated. The corresponding degrees of saturation for adult hæmoglobin are about 20% and 70%.

4. **Gaseous Interchanges in Placenta** (Fig. 675).—Direct determinations (in the cow) showed that in the mother the arterial blood was 90% saturated with oxygen, and at an O_2 pressure of 70 mm. Hg; the blood in the uterine vein was 70% saturated, and at an O_2 pressure of 41·5 mm. In the fœtus

[1] Beaven *et al.*, *Biochem. J.*, 1951, *49*, 374.

the oxygen pressure of the blood in the umbilical artery was 5·5 mm.; that in the vein was 11·5 mm. As the blood flows through the maternal sinuses the O_2 pressure falls from the arterial to the venous pressure level (*i.e.* from 70 to 41·5 mm.); as the chorionic villi dip in at many points along the maternal sinus (Fig. 672) they are exposed to an average pressure somewhere between the arterial and the venous. *Nothing like pressure equilibrium is attained* between the maternal and fœtal blood in the placenta; thus the O_2 pressure in fœtal blood leaving the placenta in the umbilical vein (11·5 mm.) is 30 mm. lower than the lowest pressure attained in the maternal sinuses. The main reason for the unsatisfactory rate of diffusion is presumably structural: in the lungs the oxygen in the air sacs is separated from the pulmonary capillary blood only by two thin endothelia (of alveoli and capillaries). The fœtal capillaries, on the other hand, are covered by much thicker and presumably much less permeable cell layers. These data make it certain that the O_2

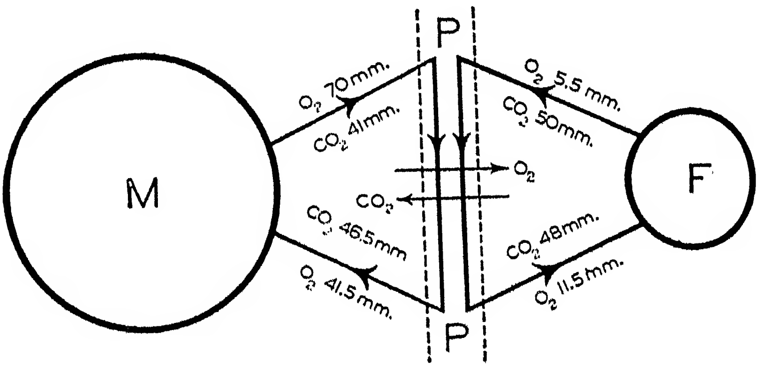


FIG. 675.—Gaseous Interchanges in the Placenta (Cow). (Roos and Romijn, *J. Physiol.*, 1938, *92*.)

M = mother; F = fœtus; P = placenta.

The arrows show the direction of flow of the maternal and fœtal circulations and the direction of passage of CO_2 and O_2.

transfer to the fœtus can be fully accounted for by simple diffusion processes. If the fœtal blood had the same oxygen-binding properties as are found after birth, very grave anoxic symptoms would develop. An O_2 pressure of 11·5 mm. (Fig. 674) would then correspond to an oxygen saturation of less than 10%; in actual fact the fœtal arterial blood is 50% saturated, owing to the change in the hæmoglobin described above. The distinctive properties of fœtal hæmoglobin thus enable it to take up large quantities of oxygen at very low oxygen pressure, and so to overcome the twin disadvantages of a poorly permeable diffusion membrane in the placenta and of comparatively low O_2 pressures in the maternal sinuses.

The CO_2 pressure values in the same experiments were: uterine artery 41 mm., uterine vein 46·5 mm.; umbilical artery 50 mm., umbilical vein 48 mm. These differences of CO_2 pressure between fœtus and mother, coupled with the much higher rate of diffusion of CO_2 (p. 366), enable CO_2 transfer to be satisfactorily effected.

THE TESTIS[1]

The testis consists of (i) *seminiferous tubules* which form the sperms and (ii) *interstitial cells* which secrete the androgen *testosterone* (Fig. 676).

Seminiferous Tubules. Spermatogenesis.—The tubules are lined by a thin basement membrane, internal to which are the *spermatogonia*, which divide to form *spermatocytes*. These in turn give rise to more medially situated smaller cells with deeply staining nuclei, the *spermatids*, and finally to *spermatozoa* (sperms). The spermatocytes undergo a reduction division by which their chromosome number is reduced by half, *i.e.* to 24 compared with 48 in somatic cells. The 24th pair (or sex pair) of chromosomes in the male consists of two distinct entities called X and Y. The reduction division produces two kinds of sperm; half contain 23 chromosomes + X, and the other half 23 similar chromosomes + Y. In the ovum, the 24th (or sex) pair of chromosomes consists of two identical entities called X and X;

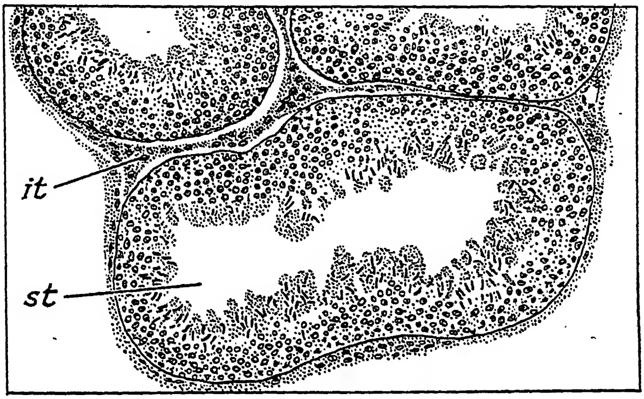


FIG. 676.—Structure of Normal Testis.
st=Seminiferous tubules showing spermatogenesis. Some interstitial cells (*it*) are seen between the tubules. (After Cowdry, *Text Book of Histology*, 1950).

after the reduction division all ova contain 23 chromosomes + X. The ovum may be fertilized by either kind of sperm. Ignoring the 23 chromosomes which are unrelated to sex, the result of fertilization may be:

Sperm X + Ovum X=Offspring XX=Female.
Sperm Y + Ovum X=Offspring XY=Male.

The sex of the offspring is thus determined exclusively by the sperm and is quite independent of the ovum (popular belief to the contrary notwithstanding).

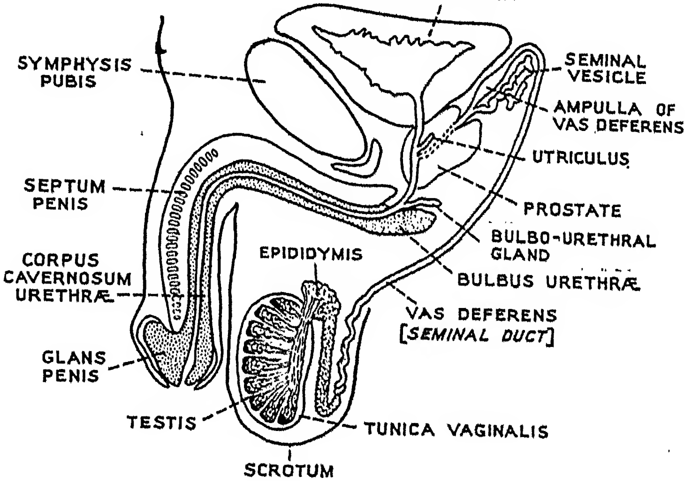
In many species spermatogenesis occurs only during a restricted breeding season. In man, too, no spermatogenesis occurs before puberty; but subsequently the process takes place continuously until it stops in elderly people. The sperms in the seminiferous tubules, in spite of their fully formed

[1] Allen, *Sex and Internal Secretions*, 2nd edn., Baltimore, 1939. Robson, *Recent Advances in Sex Physiology*, 3rd edn., London, 1947. Pincus and Thimann, *The Hormones*, N.Y., 1948, *1*; 1950, *2*. Hooker, *Recent Progress Hormone Research*, 1948, *2*, 173. Nelson and Heller, *ibid.*, 197, 228. Hamilton, *ibid.*, 257. Nelson, *ibid*, 1951, *6*, 29.

ls, are *non-motile* and are believed to be pushed passively onwards by ɘ pressure exerted by the steadily formed fresh cells into the straight bules and ductuli (vasa) efferentia. The latter channels are lined by a iated epithelium and a thin muscle coat, which propel the sperms and the in fluid in which they are suspended into the epididymis.

Between the spermatogonia are the elongated supporting *cells of Sertoli* ıich project into the lumen of the tubule ; the spermatids and sperms are tached to the inner margin of these cells and are alleged to be nourished · them.[1]

Little is known about the normal rate of sperm formation. It depends ; the integrity of the *pituitary* ; after hypophysectomy the seminiferous bules atrophy. Spermatogenesis is stimulated in the hypophysectomized



G. 677.—Anatomy of Seminal Tract and Related Glands. (After Cowdry, *Text Book of Histology*, 1950.)

ıimal by androgens (p. 1113). Frequent ejaculations lead to a progressive duction in the sperm count in the semen.

Seminal Tract and Related Glands.[2]—The seminal tract consists of e epididymis, the ductus (vas) deferens, its terminal ejaculatory duct ıich opens into the prostatic urethra, and the penile urethra. The seminal ısicles open into the ejaculatory ducts, the prostatic glands into the prostatic ethra, and the bulbo-urethral glands into the penile urethra (Fig. 677).

(1) Epididymis.—The convoluted ductus epididymis is lined by a muscle at and a secretory columnar epithelium ; it is the main storehouse of erms. While the sperms are within it they become *motile, mature* in other ıys and develop a *greater fertilizing power*. Motility is studied *in vitro* in a ırm saline solution ; fertilizing power is assayed by determining the

[1] There seems to be no evidence for this oft-repeated statement ; most cells manage nourish themselves without the aid of handmaidens.

[2] Retief, *Clin. Proc.*, 1948, *8*, 31.

incidence of pregnancy after carrying out artificial insemination with the sperms.

The fertilizing power of sperms from the distal end of the epididymis is twice as high as that of sperms from the proximal end. It seems therefore that the secretion of the epididymis normally exerts a "spermatotrophic" action. If the epididymis is tied off at both ends, the sperms retain their motility for 60 days and their fertilizing power for 30 days. If the testes are removed immediately after tying off the epididymis, both these periods are reduced by half. These effects of castration are due to absence of testicular hormone; they are overcome by injections of androgen, which restore the secretory activity of the epididymis to its normal state. If seminal ejaculation does not occur, the sperms in the epididymis ultimately degenerate and undergo liquefaction; they are not passed into the vesiculæ seminales or voided in the urine.

(2) Ductus Deferens.—This is lined by a powerful muscle coat and a columnar non-ciliated epithelium.

(3) Vesiculæ Seminales.—Each vesicle is a muscular convoluted tube lined by an epithelium which secretes an alkaline yellow viscid fluid which forms much of the volume of the ejaculated semen.

(4) Prostate.—The glands of the prostate consist of many follicle-like spaces leading into ducts. The epithelium of the follicles secretes the prostatic fluid, which is thin and opalescent and gives the semen its characteristic odour. Between the follicles there is a good deal of muscular tissue.

Prostatic Fluid.—This fluid in man is slightly *acid* in reaction (pH=6·4). It is rich in *calcium* and *citrate* (30 and 150 m.Eq./L respectively) and in *acid phosphatase.* Fluid expressed from the resting gland by digital compression contains 100–1200 units of acid phosphatase per 100 c.c.; fluid secreted during a period of sexual excitement contains 1900–4000 units (cf. p. 1114).

(5) Bulbo-urethral Glands.—These form a mucoid secretion which is discharged into the anterior (penile) urethra.

Seminal Fluid.—(i) The semen consists of the products of the seminiferous tubules, the seminal tract (especially the epididymis) and the related glands, *i.e.* the seminal vesicles, the prostate and the bulbo-urethral glands. The fluid part is contributed chiefly by the prostate and seminal vesicles.

(ii) Human semen is liquid when ejaculated, but soon coagulates *in vitro*; after 15 minutes it undergoes secondary liquefaction. Semen contains fibrinogen and thromboplastin, but no prothrombin or thrombin. Though it is rich in calcium, the excess citrate (from the prostate) must largely remove the calcium from the ionic state. The mechanism of coagulation of semen is obscure, but it presumably involves the conversion of its fibrinogen into fibrin. The secondary liquefaction of the fibrin is due to a specific enzyme, *fibrinolysin,* present in prostatic fluid. At 37° C., 2 c.c. of prostatic fluid can liquefy 100–1000 c.c. of clotted human plasma in 18 hours.

(iii) The volume of semen is 2–4 c.c. per ejaculation, which contains on an average 200 million sperms.

Male fertility is reduced clinically when (*a*) the total sperm count falls below 60 million, (*b*) more than 20% of the sperms have abnormal heads, or (*c*) the motility of the sperms in a fresh specimen of semen is impaired.[1]

(iv) Sperms form *hyaluronidase,* an enzyme which liquifies the hyaluronic

[1] Farris, *Brit. med. J.*, 1951, ii, 1476.

acid found in mucus and in the zona pellucida of the ovum. In this way the sperms can penetrate the normal plug of mucus in the cervix uteri and get through the zona pellucida which forms a tough lining surrounding the ovum.

(v) The reaction of semen is alkaline, the acid prostatic fluid being neutralized by the other components; sperms are rapidly immobilized in an acid medium.

Interstitial Cells.—These cells develop from the mesoderm of the embryo; they are abundant in the fourth month of fœtal life, fewer in the new-born, and continue to diminish to the end of childhood. There is an *increase in their number at puberty*; they remain constant in number during sexual life in man and diminish in old age. The interstitial cells are internally secreting cells and are usually arranged round the blood vessels; material with characteristic staining reactions can be traced from the cells into the capillaries. The androgen secreted by these cells is testosterone (p. 1110).

It is an interesting and unexplained fact that testis extracts also contain *œstrogen* (p. 1077). Certain testis tumours (teratomata) form *chorionic gonadotrophin* (p. 1088).

Secretion of testosterone presumably commences at about the age of puberty; androgen, however, appears in the urine earlier. In man, the primates, and the rat, testosterone is continuously secreted; most mammals, however, are seasonal breeders, and in them the secretion of the hormone is correspondingly intermittent. Secretion of the hormone is depressed by under-nutrition, and especially by vitamin-*B* deficiency.

Bodily Changes at Puberty.—At puberty the testes increase rapidly in size (Fig. 600, D) and (as stated) spermatogenesis sets in. The interstitial cells begin to secrete testosterone, and as a result the accessory organs of reproduction (epididymis, vesiculæ seminales, prostate, penis) begin to grow and the secondary male sex characters make their appearance. The scrotal skin thickens; there is growth of hair on the face, trunk, and axillæ; the pubic hair develops considerably and its upper border is convex upwards. Growth of the larynx occurs and the voice breaks. Considerable muscular development occurs. Occasional erections and discharge of seminal fluid take place.[1] Striking psychological changes also begin to make their appearance.

Physiology of Coitus.—The introduction of sperms into the vagina involves erection of the penis and ejaculation (emission) of the seminal fluid. Both processes are fundamentally reflex in character and occur quite efficiently in a spinal man following stimulation of the glans penis or related skin areas (p. 693). In the intact man any or many of the sense organs may constitute a source of appropriate afferent impulses; the response is long-circuited through the brain and involves the activity of the highest cortical levels which can modify the reaction either by way of reinforcement or inhibition. There is no need to stress the enormous importance of psychological influences and especially of emotional states on the act of intercourse. The results of castration show that the reflex arcs are influenced at some point by the internal secretion of the testes. The changes occurring in coitus are considered below.

(i) On the efferent side, *erection* is brought about by the *nervi erigentes* which relax the muscle coat of the arterioles of the penis and of the spongy

[1] Such emissions are not harmful physically and need arouse no feelings of guilt.

tissue of the corpora cavernosa and spongiosa; at the same time the dorsal vein of the penis is compressed. The penis, which in the resting state is small, flabby and covered with wrinkled skin, becomes thickened, elongated and rigid and thus well adapted for introduction into the vagina; the angle which the erect penis makes with the trunk follows closely that of the vagina and its length is such that in people of average build the semen is deposited high up in the posterior part of the vagina.

(ii) Friction between the glans penis and the vaginal mucosa, reinforced by other afferent streams and psychological factors, causes a reflex discharge along the *sympathetic to the seminal pathway*; the muscle coats of the epididymis, ductus deferens, the seminal vesicles and the prostate contract, and the sperms accompanied by the secretions of the accessory glands (p. 1103) are discharged into the posterior urethra between the internal and external sphincters of the bladder. The semen is thence ejected by the rhythmic contractions of the bulbo- and ischio-cavernosus muscles (supplied by *somatic* nerves). More prostatic fluid is also *secreted*, probably owing to *parasympathetic* stimulation of the glands. During coitus the entire urethra thus takes on a sexual function. It is important to note that the sympathetic nerves which are motor to the seminal tract also close the internal vesical sphincter and thus prevent a reflux of semen into the bladder; the contraction of the sphincter vesicæ and the associated inhibition of the detrusor vesicæ prevent a simultaneous discharge of urine.

(iii) The account given above of the innervation of the accessory reproductive organs is supported by sound clinical evidence. Stimulation of the hypogastric (sympathetic) nerves at operation produces ejaculation of semen in man. Bilateral lumbar sympathectomy below L_2 or section of the presacral nerve abolishes ejaculation, although penile erection and sensation remain normal. A lesion of all the sacral nerves below S_1 which severs the sacral parasympathetic outflow abolishes erection and produces relative anæsthesia of the penis.

(iv) The "orgasm" just described in the male should coincide in satisfactory intercourse with appropriate psychological and reflex reactions in the female, consisting in the latter of engorgement of the vulva, relaxation of the adductor muscles of the thigh and of the vaginal orifice, secretion of mucus by the vulval and vaginal glands and erection of the clitoris. The vagina becomes distensible, its lining is lubricated and it becomes easily traversable by the penis. Afferent impulses from the stimulated clitoris may heighten the state of psychical excitement in the female and help to promote a complete orgasm. It is claimed that the uterus may execute rhythmic movements which help to aspirate the seminal fluid into its lumen, but the evidence is not conclusive.[1]

(v) The physiological changes which take place in intercourse are by no means restricted to the reproductive organs and adjacent parts. The usual accompaniments of certain kinds of emotional tension are present, *e.g.*

[1] The technique, courtesies, and æsthetics of sexual intercourse are matters of outstanding importance, yet they are never taught by the physiologist and rarely discussed adequately at any stage in the medical curriculum. Sexual relations between civilized men and women are more than a matter of anatomy and physiology. The reader is advised to consult such popular works as Marie Stopes' *Married Love* and *Wise Parenthood*; Griffith, *Modern Marriage*; Fielding, *Parenthood*; Van der Velde, *Fertility and Sterility in Marriage*.

acceleration of the heart (to 150 beats per minute), rise of blood pressure, rapid breathing, flushing of the face, and sweating. Adrenaline is doubtless poured out and it is likely that the anterior pituitary (p. 931) and the thyroid glands are stimulated too.

(vi) It is not known how the sperms find their way to the ovum, which is probably lying in the Fallopian tube. Normal sperms are of course motile and can move at the rate of 1–3 mm. per minute; but the vagina and uterine cavity must represent a vast uncharted sea in which the current of fluid set up by the movements of the cilia is in an antagonistic direction, *i.e.* towards

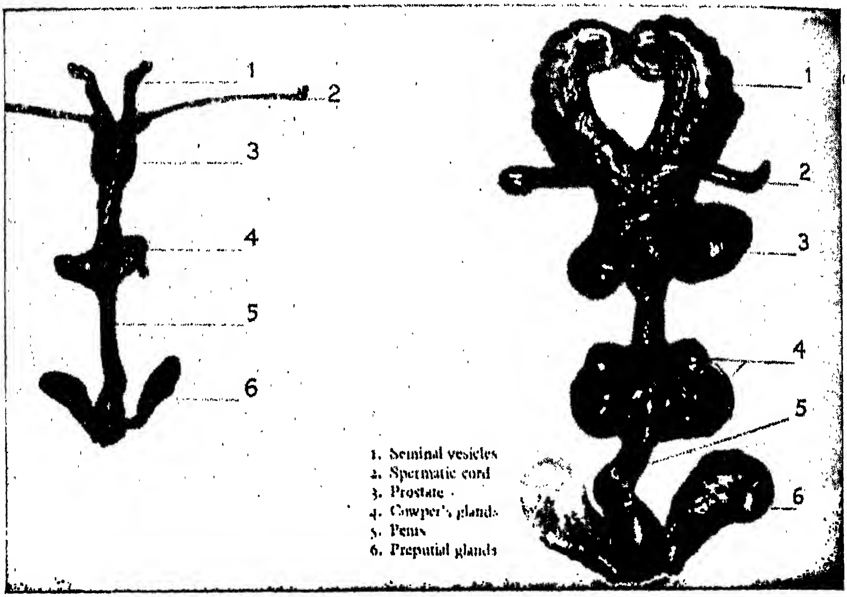


FIG. 678.—Effects of Castration and Androgen Treatment on Secondary Sexual Apparatus in Immature Rat. (Ciba's *Handbook of Internal Secretions*.)

Left.—Accessory organs of castrated animal showing marked atrophy.
Right.—Organs of castrated animal eleven days after treatment with androgen, showing complete recovery.

the exterior. There is no evidence that the ovum exerts a positive chemical attractive influence. But the 2–4 c.c. of semen discharged at one emission may contain up to 700 million sperms (mean normal=200 million); as only *one* need arrive for fertilization to be completed, blind chance may be the only factor involved.

(vii) The survival time of sperms introduced into the vagina, though a matter of great importance, has not yet been determined with certainty; it may be a day or two.[1] It is said that sperms become non-motile in the vagina in less than 60 minutes after coitus, but remain motile in the cervix and body of the uterus for 25 to 40 hours; non-motile sperms of course cannot

[1] The House of Lords has rejected the suggestion put forward in a divorce suit that there might be an interval of some 80 days between the last act of coitus and fertilization. This decision was certainly not a rash one.

move "upstream" against the current. If an orgasm occurs in the female, the sperms are said to enter the uterine cavity within three minutes, but in the absence of orgasm only after one hour.

Extirpation of Testes.—(1) BEFORE PUBERTY.—If the testes are removed there is permanent sterility; owing to absence of testosterone the usual pubertal changes do not occur and the *accessory organs of reproduction do not develop*; the vesiculæ seminales and prostate remain small and atrophic (Fig. 678). There is no growth of hair on the face, trunk, or axillæ; the pubic hair is of the female type, the outline being concave upwards; the growth of the larynx is arrested. There may be abnormal deposition of fat; accumulations may be found on the buttocks, hips, pubis, and breasts. The

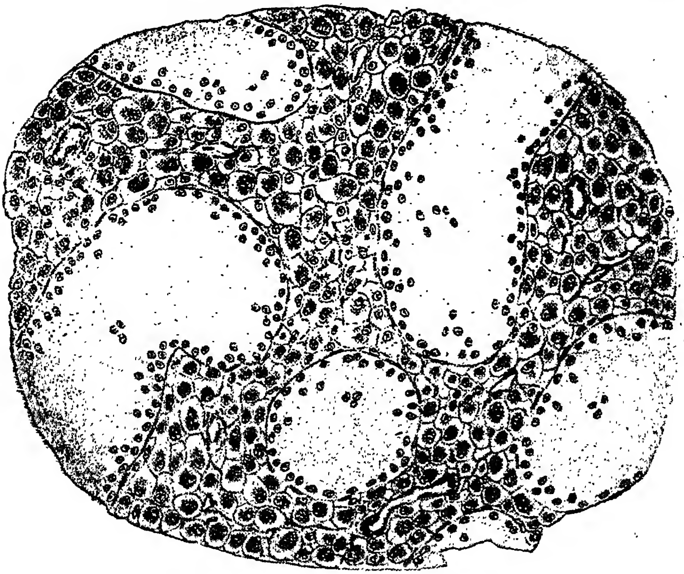


FIG. 679.—Effects on Testis of tying Vas [Ductus] Deferens.
Seminiferous tubules atrophied, interstitial cells persist as masses of cells. (Ancel and Bouin from Sharpey-Schafer, *The Endocrine Organs*.)

muscles are soft and poorly developed. There is some delay in the union of the epiphyses, but *no* regular tendency to gigantism: some eunuchs are short, others are tall, and on the whole they show the same range of variations as do normal people. The skin is pale and tans poorly when exposed to the sun.

(2) AFTER PUBERTY.—Castration after puberty produces changes which vary very much in degree in different subjects. It should be remembered that some of the secondary sexual characters and accessory organs depend on testosterone not only for their development but also for their *maintenance*; these characters or organs are depressed after castration. Thus the seminal vesicles and prostate always atrophy. Other characters or organs having once developed under the influence of testosterone can *persist in its absence*; these are unaffected by castration. There is, for example, no alteration in the voice, the penis remains of normal size (it was usually amputated in the

case of Eastern eunuchs) and the beard may be unaffected. In some cases the general bodily changes resemble those described for prepubertal castration, in others they are not very marked. Sexual desire and erection may be absent; but there are many instances in which sexual activity was little impaired, successful coitus (with ejaculation of fluid from the prostate or seminal vesicles) being frequently carried out for as long as twenty-five years after castration; in these cases the pattern of reflex behaviour which was initially induced by testosterone subsequently persisted in spite of its absence. Rats or guinea-pigs may copulate for months after castration and human eunuchs are often quite promiscuous—ten out of twenty-five studied were found to be suffering from gonorrhœa.

There is no evidence that castration damages any essential functions except those related to reproduction; it does not shorten life or produce

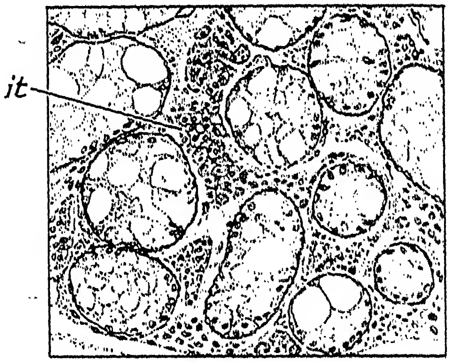


FIG. 680.—Testis in Experimental Cryptorchidism. (Allen, *Sex and Internal Secretions*, 1939.)

Section of testes (of adult guinea-pig) which had been confined in the abdomen for three months. Note that the seminiferous tubules are very degenerated. The interstitial tissue (*it*) persists.

premature senility. Many castrates have shown the highest intellectual attainments. In some, it is true that a peculiar mental state may result; this is more likely to be due to the psychological trauma produced by the castration than to loss of any internal secretion formed by the testis. If modern psychology is to be believed, the mere subconscious fear of castration may produce serious mental symptoms, although the testes are functioning perfectly normally; it is hardly surprising that the mental results of a real castration are worse. Lucius Apuleius, even after his transformation into an ass, feared castration more than death.

RESULTS OF TYING VAS DEFERENS.—If the vas [ductus] deferens is ligated, the subject becomes sterile, but no changes occur in either the somatic or the psychical sexual characteristics. The interstitial cells always persist as definite masses of epithelial cells (Fig. 679) and presumably continue to release their hormone. The changes in the seminiferous tubules are variable: sometimes they undergo complete degeneration (Fig. 679); *more commonly*

however some spermatogenesis continues, the newly formed sperms undergoing liquefaction and being replaced by freshly formed cells.[1]

CRYPTORCHIDISM.—If the testis fails to descend into the scrotum, the seminiferous tubules remain infantile in structure and no development of sperms takes place. If the condition is bilateral the individual is sterile,

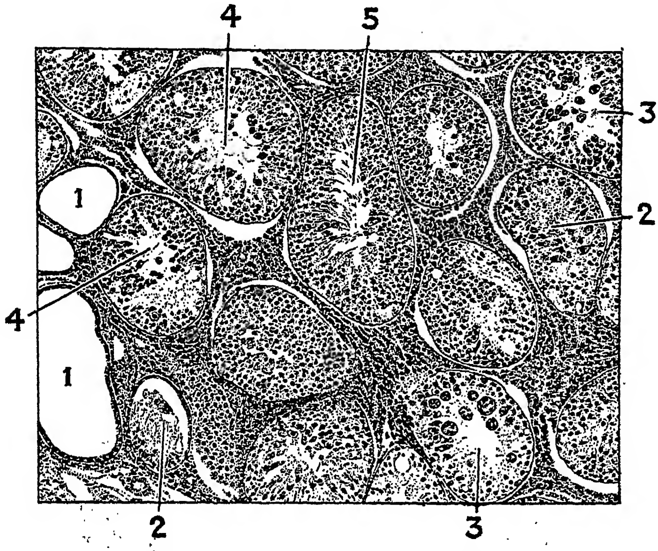


FIG. 681.—Changes in Grafted Testis. (Allen, *Sex and Internal Secretions*, 1939.)

Portion of rat testis recovered from scrotum four months after transplantation. Tubules show all degrees of change from complete destruction (1), partial destruction (2, 3, 4), to almost normal tubules containing spermatozoa (5).

but as the interstitial cells are structurally normal and continue to secrete, the secondary sexual characters develop normally. The lack of development of the seminiferous tubules in the cryptorchid is attributed to the higher temperature to which the gland is exposed in the abdomen compared with that in the scrotum ; if the scrotum is kept artificially warmed, of if the testis

[1] An interesting case is recorded of a man whose wife, believing that pregnancy would ruin her health, persuaded him to have both his vasa tied in 1939 ; in addition a short length of each vas was excised. Subsequently there was complete absence of sperms from the seminal fluid. Owing to temperamental incompatibilities the marriage was later dissolved and the man married a second wife who was anxious to have children. An operation was performed (on 30th September 1949, *i.e.* after an interval of ten years) to restore the continuity of the vasa. On 8th October the semen contained no sperms ; on 10th November it contained 9 million sperms per c.c., the percentage of motile sperms being 10% and of abnormal forms 20%. In February 1950 his wife became pregnant. Examination of the seminal fluid on 28th December 1950 showed a volume of 2·5 c.c., a total count of 80 million sperms per c.c. and 50% of motile sperms. It is clear, therefore, that the spermatogenic potentiality of the testis may persist in spite of years of obstruction of the vasa. [Handley, *Arch. Middlesex Hosp.*, 1951, *1*, 74.]

is deliberately transferred to the abdomen (Fig. 680), spermatogenesis ceases.[1]

The differences noted between the effects of cryptorchidism and those of castration demonstrate strikingly that the interstitial cells are responsible for the internal secretion which controls the accessory reproductive organs and the secondary sexual characters.

TESTICULAR GRAFTS.—Testes grafted into another animal always finally undergo atrophy (Fig. 681); it is uncertain to what extent such grafts liberate their internal secretion. Apart from these doubts, testicular hormone can be so readily and effectively administered in simpler ways that grafting as a therapeutic method is quite unjustified.

Testis Hormones. Androgens (Testoids).—The term *androgen* is used to describe any substance which has masculinizing properties, *i.e.* which promotes the growth of the accessory organs of reproduction in castrated male mammals (Figs. 678, 683), or of the comb, wattles, and ear lobes in castrated male birds (Fig. 682). The androgen secreted by the testis is *testosterone*; its chemistry and that of related substances is reviewed on

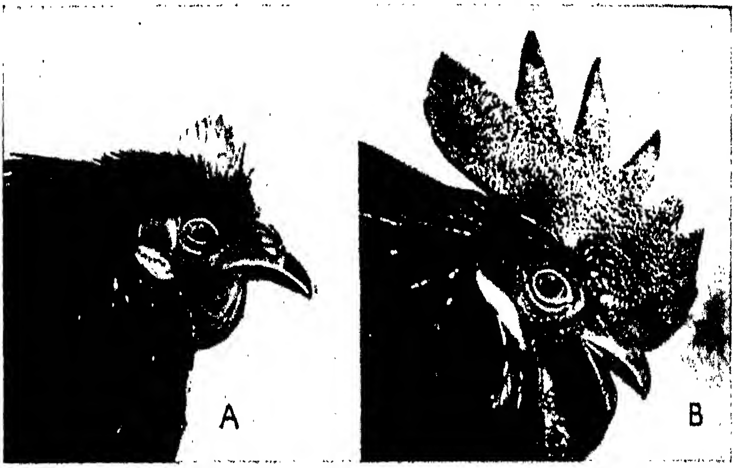


FIG. 682.—Effects of Androgen on Capon. (Koch, *Bull. N.Y. Acad. Med.*, 1938, 14.)

A. Head showing comb and wattles of castrated bird.
B. Extensive growth of comb and wattles as a result of intensive treatment for eighteen days with androgen.

p. 1076. Two weaker derivatives, *androsterone* and *isodehydroandrosterone*, are found in urine and represent degradation products of testosterone. They are neutral 17-ketosteroids and constitute a small fraction of the total excretion of these substances in male urine (p. 953). The liver is the main site of inactivation and modification of androgens.

Androgen can also be extracted from the adrenal cortex (p. 967) and is present in virilizing ovarian tumours.

[1] Inadequate attention has been paid to the effects (if any) on spermatogenesis of wearing the kilt or of residence in a hot country. If the adverse effects of raised temperature on spermatogenesis are as great as suggested in the text one would expect to find a markedly reduced fertility in the Tropics; of this there is no evidence.

Many other androgens have been artificially prepared, the most important clinically being *methyl testosterone*, which is active *by mouth*. Even more than is the case with the œstrogens the activity of testosterone is enhanced when it is combined with fatty acids ; this is especially so in the case of testosterone *propionate*.

ACTION OF ANDROGENS.—(i) Naturally secreted testosterone is responsible for the development of the accessory organs of reproduction and the other secondary male sexual characters at puberty, and in the case of some of them for their persistence throughout adult life (p. 1107). When injected into an immature animal androgen causes precocious development of the accessory sexual organs.

(ii) Androgens overcome the degenerative changes in the accessory sexual organs resulting from castration. Thus castration produces atrophy

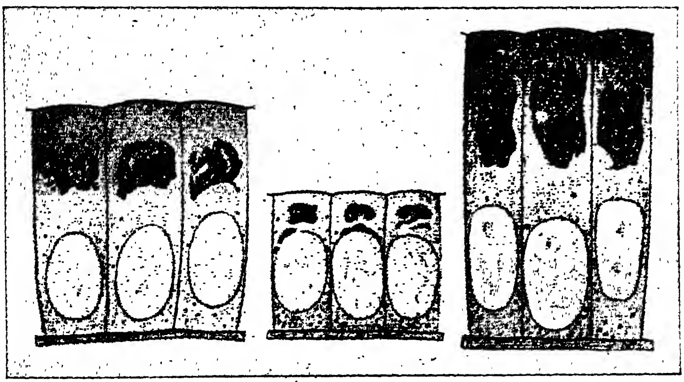


FIG. 683.—Influence of Castration and of Androgen on Cytology of Prostate. (After) Moore, Price, and Gallacher, from Allen, *Sex and Internal Secretions*, 1939.

Cells of prostate of rat showing (black) Golgi apparatus. From left to right: normal cells; cells from animal castrated twenty days previously; similar castrate treated with androgen. Note resulting increase in size of cell and recovery of Golgi apparatus.

of the prostate and degeneration of the glandular epithelium ; injections of androgen can restore the prostate to normal even after an interval of 6 months after castration (Fig. 683).

(iii) By an action on the epididymis androgen first enhances and then maintains the motility and fertilizing power of the sperms which are stored there (p. 1103).

(iv) It is doubtless responsible in large part (initially at least) for the distinctive emotional make-up of the male.

(v) *Action on Seminiferous Tubules.*—In the hypophysectomized animal administration of androgen prevents the degeneration of the seminiferous tubules which otherwise occurs, and maintains spermatogenesis (p. 1113).

Clinical Use of Androgens.—These are well illustrated by the following case reports.[1] The first is of a thin under-developed boy with the usual signs

[1] Foss, *Lancet*, 1937, ii, 1307; 1938, ii, 1284; *Brit. med. J.*, 1939, ii, 11. Spence, *Quart. J. Med.*, 1940, 9, 309.

merely by slight penile turgidity; erection did not occur. He was given 20 mg. (daily) of testosterone propionate by injection. From the first day of treatment sexual desire was increased, erection occurred readily and was accompanied by good erotic sensation; coitus took place nightly, though no ejaculation took place (doubtless from lack of adequate secretion by the prostate and seminal vesicles). On the eighth day of treatment he was unable to sleep all night on account of persistent and painful erection, which was unrelieved by coitus. The erect penis now measured 15 cm. in length, 11 cm. in circumference at the glans, and 11·5 cm. at the base; it is reported that he now "fully satisfied his wife." He gained 16 lb. in 12 weeks and had to wear bigger collars. After an interval he was treated with equal success, with testosterone propionate solutions rubbed into the skin and with this course ejaculation successfully occurred. It is interesting to note that hair grew on the regions of thigh, calf, and leg to which the hormone was applied.

Fig. 684 illustrates the marked growth of the pubic hair and enlargement of the penis which were produced in an adult patient with defective testicular internal secretion as a result of three months' treatment with testosterone propionate.[1]

Control of Testicular Activity.—(1) ANTERIOR PITUITARY.—The anterior pituitary by means of its gonadotrophic hormones controls the growth, functional integrity, and activities of the testis. The exact mechanisms involved are still uncertain and are discussed below.

(i) Hypophysectomy in immature male animals causes the testes to remain infantile; the accessory reproductive organs do not develop owing to lack of testosterone. Hypophysectomy in adults leads to testicular atrophy and the same changes in the accessory reproductive organs and elsewhere as follow castration. In Simmonds' disease, hypogonadism (depressed spermatogenesis and decreased testosterone secretion) is a common finding. Anterior pituitary grafts made into immature male animals hasten the onset of testicular maturity and of the other signs of puberty; in hypophysectomized adult animals they restore normal testicular function. Anterior pituitary grafts made into normal animals may stimulate excessive secretion of testosterone and hence cause overgrowth of the accessory reproductive organs, especially of the vesiculæ seminales.

(ii) *The interstitial-cell-stimulating hormone* (*ICSH*) initiates and sustains the internal secretory activity of the interstitial cells. ICSH is identical chemically with the luteinizing hormone (LH) of the female. The secretion of testosterone secondarily causes growth and development of the accessory organs of reproduction. If ICSH is injected into hypophysectomized animals it also causes the return of *spermatogenesis*; this action is *not* a direct one on the seminiferous tubules but is due to the release of testosterone which in its turn acts on the tubules. This interpretation is supported by the facts that in hypophysectomized animals and in patients with Simmonds' disease the administration of androgen restores spermatogenesis. In the *intact* animal, however, the action of administered androgen on the testis is small and variable.

(iii) It has been supposed that the anterior pituitary (via the hormone known as FSH in the female) directly controls spermatogenesis. This hormone

[1] The action of androgen in the female and its possible clinical uses in women is considered on p. 1082.

was therefore called "gametokinetic" to indicate that in both sexes it regulates the formation and maturation of the gametes (ovum or sperm). If pure FSH is injected into hypophysectomized animals it does *not* induce spermatogenesis. If it is given together with ICSH it is claimed that the restoration of spermatogenesis is more complete than with ICSH alone. FSH may thus be an *accessory direct* stimulating factor on the tubules and may be necessary for the maintenance of an optimal level of spermatogenetic activity.

(iv) The testis in its turn influences the activity of the anterior pituitary. After castration in animals the basophil cells in the pituitary increase in size and number and there is increased secretion of gonadotrophin. The testis thus normally inhibits the secretion of gonadotrophin. In clinical eunuchoidism due to *primary testicular failure* there is also excessive gonadotrophin formation, which is relatively immune to androgen treatment but is readily inhibited by œstrogens. (When eunuchoidism is *secondary to pituitary insufficiency* there is of course decreased secretion of gonadotrophin.)

The source and nature of the testicular inhibitory factor are still under consideration: (*a*) it might be androgen; but as mentioned above excess gonadotrophin secretion is *not* inhibited by physiological doses of androgen; (*b*) it might be some other unidentified testicular "factor," derived perhaps from the seminiferous tubules.

(v) *Klinefelter Syndrome.*[1]—Light is thrown on this problem by the findings in this uncommon syndrome which has the following characteristic features: (i) the seminiferous tubules are severely damaged or hyalinized; the germinal epithelium and Sertoli cells are absent and no sperm formation occurs; (ii) the interstital cells are normal in appearance and secretory activity as judged by the normal state of the accessory reproductive organs; (iii) there is considerable enlargement of the breasts owing to duct proliferation and increase in connective tissue; (iv) there is increased excretion of gonadotrophin (FSH) in urine; (v) the urinary output of 17-ketosteroids is reduced.

The excessive output of gonadotrophin suggests that the pituitary has been released from a normal inhibitory control. The testicular lesion however affects the tubules only; androgen secretion is unaffected. It is, therefore, concluded that the seminiferous tubules normally secrete into the blood an unidentified pituitary-inhibiting factor tentatively called "inhibin."

The possible inter-relationships of anterior pituitary and testis are shown in Fig. 685. It is probably unwise in the present state of knowledge to press too closely analogies between the control of the gonads in the male and female.

(vi) The anterior pituitary may control the *descent of the testis*; in some cases of undescended testes the testes have entered the scrotum following (and perhaps as a result of) treatment with gonadotrophic hormone.

(2) The *thymus* normally persists till puberty, when the reproductive organs develop; castration prolongs the period of persistence of the thymus. The sex glands therefore exert a depressant effect on the thymus (p. 1016).

(3) For action of *thyroid* and *adrenal cortex*, see pp. 986, 967.

The Prostate.[2]—The structure, function, and hormonal control of the prostate were briefly considered on pp. 1103, 1111. Certain other aspects of the gland which are of clinical importance will now be discussed.

PHOSPHATASE IN PROSTATE.—(i) The prostate contains small and

[1] Klinefelter *et al.*, *J. clin. Endocrin.*, 1942, *2*, 615.
[2] Huggins, *Harvey Lectures*, 1946–47, *42*, 148.

unimportant amounts of *alkaline* phosphatase (0·5–1·3 units/g.) in the capillary walls.

(ii) The *acid* phosphatase content on the contrary is uniquely large; this enzyme is found in the epithelial cells and in the lumen of the glands. In man, acid phosphatase is found in the prostate in the following concentration in units/g.: new-born child 1·5; adolescent 70; adult 500–2,300. The increase in acid phosphatase content at puberty is due to the action of testosterone; thus if a prepubertal monkey is treated with androgen, the acid phosphatase in its prostate increases several hundred times.

(iii) The normal *serum* acid phosphatase content is only 3 units (or less) per 100 c.c.; it is derived from many organs, the prostate being a quite *unimportant* source because the enzyme cannot get through the walls of the prostatic vessels into the circulation; the *serum* acid phosphatase content is thus approximately the same in children and women as in adult men.

(iv) The physiological rôle of the prostatic acid phosphatase is unknown;

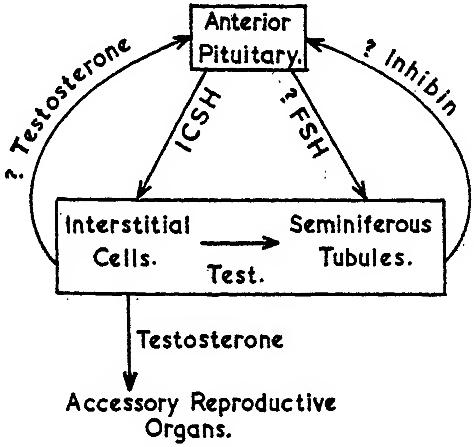


FIG. 685.—Testis–Anterior Pituitary Interrelationships.

if a suitable substrate (like hexosephosphate) were present in the gland with consequent release of inorganic phosphate, large amounts of $Ca_3(PO_4)_2$ would be precipitated owing to the rich Ca content of prostatic fluid. Calcareous masses of $Ca_3(PO_4)_2$ commonly occur in the prostate; the surprising thing is that they are not larger and more frequent.

(v) Adult male urine contains 3–5 times as much acid phosphatase as the urine of women and children; the excess enzyme in the male is derived from the prostatic *fluid*, the composition of which was described on p. 1103.

(vi) If acid phosphatase is injected intravenously it disappears from the blood in 3–6 hours; the normal level of *serum* acid phosphatase thus depends on a nice balance between its discharge into the blood from many organs and its removal from the blood.

Carcinoma of Prostate.—The cells of the carcinoma form acid phosphatase to a varying degree; on an average the enzyme content of the cancer is

1/20th that of normal prostate (the range is 19–280 units/g.). It was explained on p. 1107 that the prostate is dependent on testosterone not only for its development but also for its *maintenance* in adult life. The cells of the carcinoma of the prostate also depend on testosterone for their maintenance but the degree of dependence varies considerably with the individual tumour; (the *cause* of the carcinoma is unknown). When the growth metastasizes the secondary deposits, like the parent growth, also form acid phosphatase. When the deposits occur in the bone marrow and lymph glands the enzyme can pass *fairly readily into the blood stream* with the result that the *serum acid* phosphatase level *rises* from 3 units/100 c.c. to 10–700 units. A serum value exceeding 10 units/100 c.c. is diagnostic of *metastasis* of prostatic carcinoma. Secondary deposits in bone for some unknown

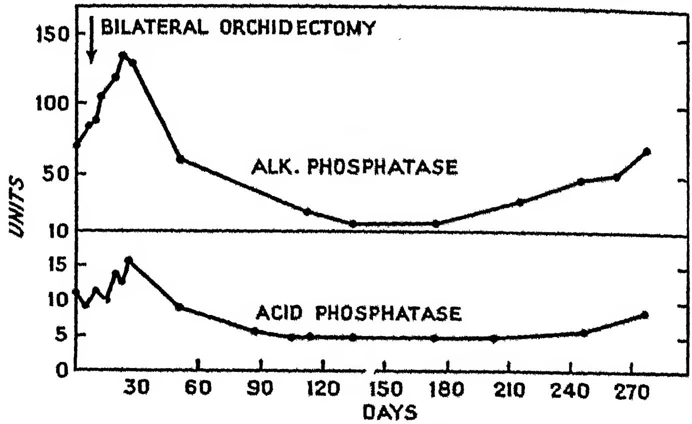


FIG. 686.—Effects of Castration on Serum Alkali Phosphatase and Serum Acid Phosphatase in case of Carcinoma of Prostate. (Huggins, *Harvey Lectures*, 1947, *42*, 188.)

Ordinate: Serum phosphatase in units per 100 c.c.
Abscissa: Time in days.
At arrow: Both testes removed.
Serum Alkaline Phosphatase. (Normal in adults 1–4 units/100 cc.). Note that the initial level is 75–100 units. A temporary further rise followed castration; subsequently a rapid decline set in. When clinical relapse occurred the alkaline phosphatase rose once more. The fall was accompanied by clinical remission, the subsequent rise (at 270 days) by clinical relapse.
Serum Acid Phosphatase.—After orchidectomy the level gradually fell from 10–15 units to about 5 units/100 cc. It began to rise when clinical relapse occurred.

reason stimulate osteoblastic activity and so raise the local concentration of *alkaline* phosphatase with a corresponding rise of the level of *serum alkaline* phosphatase, *e.g.* to 70 units/100 c.c. (Fig. 686).

EFFECT OF CASTRATION AND ŒSTROGENS.—According to the degree to which the malignant cells are dependent on testosterone, the growth (both primary and secondary) regresses after castration or on administering an œstrogen (*e.g.* stilbœstrol). Adenocarcinoma is much more sensitive to testosterone than undifferentiated carcinoma and is therefore much more susceptible to œstrogen therapy. The œstrogen acts (i) by inhibiting the release of gonadotrophin by the anterior pituitary and so decreasing natural testosterone secretion and (ii) by peripherally competing with and antagonizing

the action of testosterone. In successful cases there is clinical improvement, and a fall of serum acid phosphatase sets in after a variable period. The dwindling of the bone metastases leads to a temporary unexplained further rise of bone and serum alkaline phosphatase, which after 1–3 months rapidly declines towards normal (Fig. 686). Some patients have remained well for as long as 6 years. The administration of androgen makes the patient worse and raises the serum acid phosphatase. Generally the dependence of the tumour on androgen is partial, so that after a variable period of œstrogen treatment the growth of the tumour is resumed. Undifferentiated prostatic carcinomas are completely independent of androgen and are therefore unaffected by castration.

Action of Œstrogen on Normal Prostate.—Though the prostate has a uniform histological appearance, the effects on it of injections of œstrogen suggest that it is a dual structure. In the dog the prostate consists of a dorsal and a ventral segment; both parts respond to castration by atrophy of the glandular epithelium and in both parts the epithelium becomes tall and columnar after injection of androgen. On injecting œstrogen, however, in the *ventral* part the epithelium atrophies and the acini collapse, while in the *dorsal* part the columnar epithelium becomes stratified squamous. It has been suggested that the *posterior* lobe of the prostate in man is functionally distinct from the other lobes; it is the common site of carcinoma and never the site of the "benign" enlargement.

Prostatic enlargement commonly occurs in elderly men; the cause is unknown. It has been attributed to a decrease in androgen and an increase in œstrogen formation. Androgen therapy however does not benefit the patients.

APPENDIX

CRITICAL ANALYSES OF FIGURES

FIRST ANALYSIS

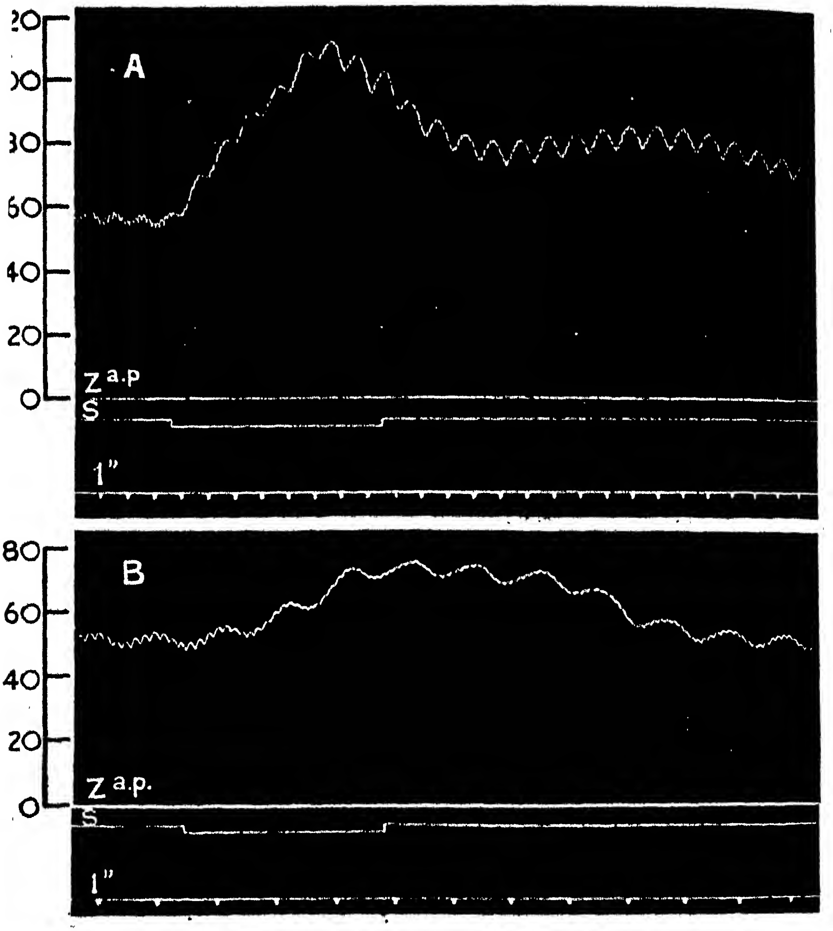


FIG. 687.—Unipolar excitation of exposed transverse surface of lateral column of spinal cord at first cervical level; effect on arterial pressure. Decapitate preparation (E. B. Cox and N. B. Dreyer).

Z a.p., zero level of pressure; S, signal marking time of faradization; 1″, time in secs. In B the quicker drum makes the cardiac acceleration more obvious.

(Figure and legend from Liddell and Sherrington, *Mammalian Physiology*, Clarendon Press, 1929.) The scales on the left represent mm. Hg.

The preparation, as described, is decapitate; the spinal cord has been divided at its junction with the medulla and the brain (and head) removed, leaving the upper surface of the first cervical segment exposed. The blood pressure record is of the kind obtained by inserting a cannula filled with some anticoagulant fluid into a large artery (*e.g.* carotid or femoral) and connecting it with a mercury manometer (p. 301). The movements of the mercury are transmitted by means of a float equipped with a writing point to the blackened surface of the paper fixed on the moving drum. As stated, the drum is moving more rapidly (about twice as fast) in record B as in A.

Control Period. The arterial blood pressure in A is fluctuating between 55 mm. Hg at minimum and 60 mm. Hg at maximum; it is a little lower in B. The arterial blood pressure in a normal (anæsthetized) or decerebrate cat is of the same order of magnitude as in an intact man. Why then is the initial level of blood pressure in this experiment so low? The blood pressure depends on the product of the cardiac output and the peripheral resistance (p. 302). The latter depends on the integrity of the vasomotor centre (p. 303) and of the sympathetic nervous system. A trans-section at the upper border of the spinal cord cuts off the tonic discharge which normally passes the vasomotor centre to the sympathetic connector cells in segments Th 1 to L 2 of the spinal cord. There is consequently universal vasodilatation owing to loss of arteriolar tone and the blood pressure falls to the level shown in the record (p. 303).

The control strip of blood pressure record shows two kinds of oscillations (better seen in the fast record B than in A). The small excursions are obviously due to the heart beat; the larger resemble the variations that commonly occur with the phases of respiration.

The control heart rate in A is 24 beats in 4 sec. and in B 14 beats per 2 sec. (corresponding to a rate of 360 and 420 per minute). This is a far higher rate than is found in the intact animal. In this preparation the vagus nerves have been cut off from their nuclei of origin in the medulla. After bilateral vagal section, normal vagal tone (or vagal restraint) is abolished; the heart rate is consequently markedly increased. (In man, following full doses of atropine to "block" the vagal terminals, the heart rate rises to about 150 per minute; the maximal heart rate in man in excitement or exertion is under 200 per minute; the rate of the released heart in this cat is far more rapid.) The extent of the oscillation of blood pressure with each beat represents the difference between systolic and diastolic pressure, *i.e.* the pulse pressure. In this experiment the pulse pressure as recorded is about 2 mm. Hg. Is the record a reliable measure of pulse pressure? The answer is that it is not. A mercury manometer possesses considerable inertia; the pressure as recorded neither rises to the full systolic nor falls to the full diastolic level, but oscillates slightly round the mean blood pressure. In this experiment, owing to the exceedingly rapid heart rate, the output per beat must be correspondingly reduced, so that the true (as well as the apparent) pulse pressure is greatly diminished.

The larger blood pressure variations, labelled tentatively "respiratory," occur at the rate of 1 per sec. or 60 per minute. But how can a decapitate preparation breathe in the absence of the respiratory centres in the pons and medulla? and how can a spinal animal survive if it cannot breathe? Presumably the animal is being kept alive by means of artificial respiration

(generally carried out with a pump connected with a cannula inserted into the trachea). The rate of artificial respiration (60 per minute) is high and it is not unlikely that the preparation is being over-ventilated and so suffering from carbon dioxide lack. Why should the blood pressure vary with the phases of artificial respiration ? The factors coming into play during natural respiration are analyzed on p. 313 ; but normal breathing differs in many respects from pump ventilation. There is, for example, no *active* enlargement of the chest cavity and no fall of intrapleural pressure during the phase of pump inspiration. Nor are there any variations in this experiment in the rate of the heart with the phases of respiration (sinus arrhythmia depends on the integrity of the vagi ; in this preparation the vagus nuclei in the medulla have been removed). During the period of inflation of the lungs by the pump, the intra-pulmonary pressure is raised as the inflated lungs forcibly expand the chest cavity ; during deflation the intra-pulmonary pressure falls. During inflation the capillaries, venules, and veins in the lungs and in the chest cavity are compressed, impeding the return of blood to the left heart and so decreasing the cardiac output and blood pressure ; the reverse occurs during deflation. If this explanation is correct then a simultaneous record of blood pressure and " respiration " would show the blood pressure falling with " inspiration " and rising with " expiration."

Period of Stimulation. During the descent of the signal line, S, the spinal cord is subjected to " unipolar excitation " and " faradization." What is " unipolar excitation " ? Two electrodes are employed as follows: one electrode is a large indifferent electrode consisting, *e.g.* of a large copper plate wrapped round with cotton wool soaked in saline and fixed firmly over hairless skin (usually the foot pad) ; the *active* electrode is a fine wire which can be applied accurately to a very restricted region which is to be stimulated. Faradization means that the primary circuit of the induction coil employed is rhythmically interrupted. The electrodes are attached to the secondary circuit probably through a short-circuiting key. " Unipolar faradization is particularly suitable for stimulating nervous surfaces where the direction of the fibres to be excited lies mainly at right angles to the plane of the surface. It is therefore fitted for experiments on point-to-point stimulations of cut planes of the central nervous system or the natural surfaces of the bulb or cortex cerebri." Stimulation is carried out in A for 8 sec. and in B for 3·5 sec. In A after a latent period of about 0·5 sec. the blood pressure begins to ascend smoothly but shows an increase in the rate of ascent towards the end of stimulation ; the peak pressure attained is 110 mm. Hg ; the pressure falls to 90–100 mm. Hg at the end of stimulation. The heart rate obviously increases but accurate measurement is impossible in A. In B the blood pressure also begins to rise after a latency of about 0·5 sec. and continues to rise throughout stimulation to a maximum of 70 mm. Hg. The cardiac acceleration sets in somewhat later (after 1 sec.), the rate rising from about 7 per sec. to about 10 or 11 per sec. (*i.e.* to over 600 per minute).

Recovery Period. On cessation of stimulation in A the blood pressure progressively declines to about 70 mm. after 4–5 secs. and stays at about this level with small fluctuations to the end of the record. The size of the blood pressure variation with " respiration " becomes more marked towards the end of the period of stimulation and remains more marked to the end of the record. In B the blood pressure returns to normal in 5–6 secs. ; the cardiac

acceleration persists after the end of stimulation, and at the end of the record the rate is still 8–9 per sec. (over 500 per minute).

Comment. The net result of the period of stimulation is a rise of arterial blood pressure and acceleration of the heart, which continue for varying periods of time after stimulation is discontinued. The simplest explanation of the results is that stimulation of the spinal cord leads to a discharge along the sympathetic nerves to the heart and blood vessels. Fig. 437 shows the chief tracts which are found in the lateral columns of the spinal cord. The fibres involved in the reaction are presumably fibres descending from the brain stem to end in the lateral horns of grey matter in the thoracic and upper lumbar cord where the sympathetic connector cells are situated. The cardiac acceleration is presumably due to stimulation of fibres which descend from the cardio-accelerator centre in the medulla to end in Th 3, 4 where the cardiac sympathetic fibres arise. The detailed route of the sympathetic fibres to the heart is described on p. 709. The rise of arterial blood pressure may be due to increased cardiac output or increased peripheral resistance or both. Sympathetic stimulation increases the force as well as the rate of the heart beat; but unless the venous return is simultaneously increased or there is a high initial level of venous pressure the cardiac output is not increased in consequence. In other words, the increased heart rate diminishes the output per beat and generally does not increase the output per minute. The tracings give no direct evidence on these points. The apparent pulse pressure is decreased when the heart is quickened, suggesting a decrease in stroke volume; but it must be remembered that, owing to the inertia of the mercury in the manometer, the size of the excursion per beat decreases when the heart quickens, even if the stroke volume is unaffected. In B it is clear that the rise in blood pressure precedes the outset of cardiac acceleration; the blood pressure also declines in B while the cardiac acceleration persists. It seems therefore that the rise of blood pressure and the cardiac acceleration are partly at any rate independent of one another. It may be assumed therefore that the rise of blood pressure is partly (if not wholly) due to increased peripheral resistance. Stimulation of the spinal cord excites the fibres passing down from the vasomotor centre to the sympathetic connector cells in Th 1 to L 2; impulses flow out along the sympathetic nerves (for route employed cf. pp. 709 *et seq.*) to the arterioles (and perhaps the capillaries and venules), produce vasoconstriction and so raise the blood pressure. Additional data would be needed to provide complete proof, such as those provided by plethysmography (p. 304), blood flow (p. 305), or temperature measurements (p. 305).

The after-effects must now be dealt with, *i.e.* the persistent acceleration in B and the persistent elevation of the arterial blood pressure in A. It is known that sympathetic stimulation involves two chemical intermediaries: acetylcholine in the ganglia and adrenaline at the sympathetic terminals in the tissues. If the acetylcholine formed during stimulation in the ganglia were to persist locally after cessation of stimulation it might continue to excite the ganglia and set up peripherally passing impulses. Similarly the adrenaline liberated peripherally might persist, so maintaining the effects of sympathetic stimulation after impulses had ceased to arrive at the blood vessels and heart muscle. There is another possibility. If there is an outflow of impulses along the sympathetic generally in this experiment then the adrenal medulla may be stimulated (p. 730) causing a secretion of adrenaline

into the circulation; this adrenaline will continue to produce peripheral effects (identical with those of sympathetic stimulation) after stimulation has been discontinued. Secretion of adrenaline during stimulation may account for the step-like ascent in blood pressure already noted in A.

As previously mentioned the extent of the blood pressure excursions which were attributed to respiration is increased in A during the latter part of the period of stimulation and very obviously after stimulation has stopped. It suggests at first sight that there is an increase in the depth of respiration. But no natural respiration is taking place and there is no reason to suppose that the stroke of the pump maintaining artificial respiration has been increased. The following suggestion may be made. The degree of expansion of the lungs depends not only on the stroke of the pump but also on the resistance encountered in the respiratory passages, especially in the muscular bronchioles. When the bronchioles are constricted, the lungs become less expanded; when the bronchioles dilate, the lungs undergo greater expansion. It is possible that in A bronchiolar dilatation occurs. It is known that sympathetic stimulation can produce broncho-dilatation.

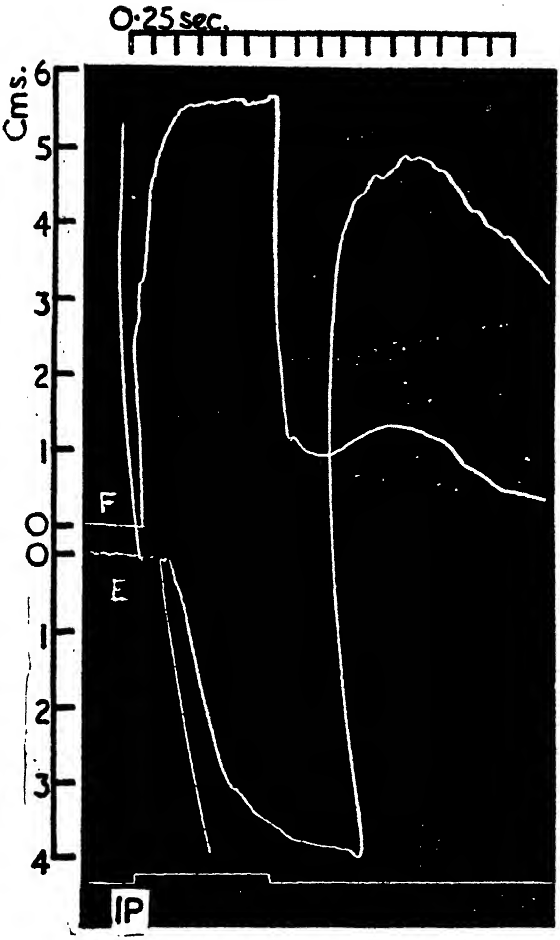


FIG. 688.—Reciprocal reflex of antagonistic muscles of knee.

F, semitendinosus, a knee flexor; E, vastocrureus, a knee extensor; IP, ipsilateral popliteal nerve. The contraction of the flexor muscle is accompanied by inhibition of the extensor; the latter is followed by a marked rebound contraction. Decerebrate cat. Time marked above in 0·25 seconds. The signal line signals upwards. The myograph writer for extensor muscle is set a little to right of that for flexor muscle in order that the two may clear each other: the ascent of F and the descent of E are therefore, in fact, practically synchronous.

(Figure and legend (slightly modified) from Sherrington, *Quart. J. exp. Physiol.* 1913 *6*, 257.)

The preparation is decerebrate, *i.e.* the brain stem has been trans-sected through or below the red nucleus and above Deiters' nucleus, generally between the superior and inferior colliculi. Such a preparation can breathe naturally and has intact central and reflex mechanisms for the regulation of heart rate and blood pressure; as the heat regulating centres have been removed, the preparation is poikilothermic. No details are given of the type of myograph employed. Muscular contraction can be studied under isometric or isotonic conditions (muscle tension or length respectively being recorded). In this example the contraction appears to be isotonic and the degree of shortening or lengthening of the muscle fibres is being studied (a cm. scale has been appended for convenience); an ascent of the lever represents contraction, a descent relaxation. Inspection of the extensor record E shows that on stimulation of IP (ipsilateral popliteal nerve) the lever falls markedly. Now completely unstimulated skeletal muscle is fully relaxed; there are no efferent inhibitory nerves to vertebrate skeletal muscle. Therefore if the extensor muscle E at some stage undergoes relaxation beyond its original control level, it must have been initially in a state of contraction. This contraction can only be due to impulses reaching E along the motor fibres from controlling ventral horn cells; as ventral horn cells cannot discharge spontaneously they must have been undergoing afferent, *i.e.* reflex stimulation. The extensor muscle initially is therefore in a state of *reflex contraction.* This is to be expected in a decerebrate preparation in which all the extensor (antigravity) muscles are reflexly contracted as a result of impulses set up by *stretch* of the antigravity muscles themselves; this proprioceptive reflex contraction represents muscle *tone* which is present in heightened form and in unusual distribution in the decerebrate preparation and subserves its imperfect standing posture (cf. p. 583). There is no evidence in the record to indicate whether the flexor muscle F is initially in a state of tone or not; generally but not invariably the antagonistic flexors in decerebrate rigidity are reciprocally inhibited.

During the ascent of the signal line IP, the ipsilateral popliteal nerve was stimulated for 1·5 sec. Ipsilateral=of the same side; *i.e.* the popliteal nerve the central end of which was stimulated (by interrupted induced current=faradization) was on the same side (in the same limb) as the responding muscles.

Results of Stimulation. Arcs have been inscribed by the two levers on the drum while stationary to indicate their direction of movement. They trans-sect the control line for F and E at the point of *onset* of stimulation.

After a latent period of about 0·1 sec., the flexor F begins to contract. The ascent shows three or four steps and reaches its summit (5·5 cm.) after 0·6 sec.; it rises very little further (with slight irregularities) to the end of stimulation. On cessation of stimulation the lever remains at its peak level or rises very slightly for a further period about equal to the latent period (about 0·1 sec.). Rapid relaxation sets in to 1·1 cm., the lever falling along an arc almost parallel to that inscribed on the stationary drum. After a trivial secondary recovery there is a further fall to 0·9 cm., followed by a secondary rise to a peak of 1·4 cm. (1·2 sec. after commencement of relaxation), with further gradual, irregular relaxation almost to base line (about 3 secs. after cessation of stimulation).

In E, after a similar latent period (0·1 sec.) the lever descends (the muscle

relaxes) first rapidly (showing one notch) to 3·2 cm. in 0·2 sec.; relaxation proceeds increasingly gradually (showing two further notches) right to the end of stimulation; the further fall is under 0·8 cm. in just over 1·0 sec. On cessation of stimulation there is a small but unmistakable after-fall (*i.e.* further relaxation) followed by a sharp and striking ascent of just over 9·0 cm. (more than 5 cm. above control level), reaching its peak in 1·5 sec.; the final part of the ascent is slower and irregular. It is followed by a slow irregular descent, but at the end of the record the level of E is still 3·5 cm. above the control level.

Comment: F. The latent period (p. 532) is due (as in all reflexes) to the time spent in conduction of the impulse in the afferent and efferent nerves, transmission at the neuromuscular junction (p. 510) and the latent period of the muscle itself (peripheral delay), plus the time spent in transmission in the central nervous system (central delay). The rate of development of contraction, *i.e.* the rate of ascent of the lever requires careful consideration. A reflex contraction is always a tetanus (never a twitch). A record of the motor tetanus of F (from faradic stimulation of its motor nerve) would have been valuable for purposes of comparison (cf. Fig. 336). Some indirect data are provided by the record. The rate of contraction of the muscle in a motor tetanus is of the same order as the rate of relaxation. It is quite clear that the rate of contraction of F is far slower than the initial substantial phase of its relaxation. The arguments set out on p. 538 indicate that the slow ascent may be due to progressive recruitment of motor neurones in the motor pool of F under the influence of continued afferent stimulation; the steps on the ascent would represent the recruitment of large additional groups of motor neurones. Reflex recruitment is the result of central summation (p. 538). Recruitment is generally less well marked in the flexor than in the crossed extensor reflex (the former response tends to be of the *d'emblée* type, *i.e.* all the motor neurones are activated approximately simultaneously). The stimulation plateau represents the completion of the recruitment process. On cessation of stimulation the tension of F is maintained at plateau level for a time equal in duration to the latent period; this post-stimulation plateau is thus the result of impulses which were still travelling in the reflex arc when stimulation stopped. Relaxation (to the extent of 5·7 cm.) then occurs immediately and is due to cessation of the discharge from the corresponding motor neurones which are no longer reflexly stimulated. The subsequent terminal delayed relaxation is not due to any peculiar inherent property of the peripheral nerve-muscle mechanism; it must represent after-discharge, *i.e.* some motor neurones continue to discharge although impulses by the direct (shortest) reflex route no longer arrive at the centre. As explained on p. 539 this after-discharge may be due to impulses travelling in long circuits in the central nervous system. The fluctuating character of the after-discharge contraction cannot be accounted for by the data available in the record; afferent impulses are, however, coming up from E, so that its rebound contraction (*infra*) may be playing some part.

E. The latent period of the reflex response of E is due partly to peripheral and partly to central delay (as for F). Relaxation of E when it sets in is far slower than the main phase of relaxation of F. The relaxation must be due to central inhibition; the afferent impulses in IP inhibit or arrest the discharge of the motor neurones of E. But as we have seen, the rate of relaxation

is slower than would be expected from simple cessation of all motor impulses to E (*e.g.* as on cessation of a peripheral motor tetanus). It is probable that the gradual relaxation which becomes progressively less rapid as stimulation continues, is due to inhibitory recruitment (cf. Fig. 340, p. 542), *i.e.* an increasing number of motor neurones in the motor pool for E are brought to rest as afferent stimulation continues ; the notches in the curve may indicate inhibition involving new large groups of motor neurones. Inhibitory recruitment is the result of central (inhibitory) summation. The after-fall (after cessation of stimulation) may be accounted for as follows. IP contains many fibres, not *all* of which are necessarily inhibitory to E ; most of them are, as seen by the results of stimulation of IP, but some may be excitatory. The response of E would depend on the algebraic sum of the central effects produced by the inhibitory and excitatory afferents in IP on the motor pool of E. These opposing central effects may not persist for the same length of time on cessation of stimulation. If it is supposed that central excitation dies down more quickly than competing central inhibition, there would be a temporary accentuation of the effective central inhibition exerted on the motor pool of E. The after-fall is followed by a very striking rebound contraction, the exact causation of which is unknown apart from the fact that it follows on central inhibition. During this period the discharge of the motor neurones of E is temporarily enhanced above that producing the initial level of tone.

With reference to F, the ipsilateral popliteal nerve IP is an excitatory afferent ; with reference to E, IP is an inhibitory afferent. There is no evidence that two kinds of fibres are present in IP, one exclusively excitatory to F, the other exclusively inhibitory to E. It is known that if the flexor F and the extensor E on the contralateral side were recorded, IP would prove to be excitatory to E (crossed extensor reflex) and inhibitory to F. The antagonistic effects of IP on ipsilateral F and E depend therefore on differences in the central terminals. Fig. 345 shows the supposed arrangement ; some collaterals from IP go to F, others to E. The synaptic terminals on the centre for F are excitatory (excitatory transmission occurring), those on the centre for E are inhibitory (inhibitory transmission occurring). The result is contraction of F (the protagonist) and reciprocal inhibition of E (the antagonist) permitting smooth flexion of the knee to occur. The tracing illustrates therefore reciprocal innervation. " In this co-ordination the inhibition is not peripheral but central, that is, it has its seat not in the muscle nor in the peripheral nerves, but in the nervous centre, probably about the starting-point of the ' final common path.' The muscle relaxes because the motor discharge from that centre is abated."

The tracing may arouse reflections about inhibition in relation to autonomic and to skeletal muscle respectively. The most appropriate comment can be given in a quotation from Sherrington :

" Nerve makes its first appearance phylogenetically in association with muscle. Of all tissues muscle is the closest and most delicate exponent of nerve-action. Yet not all muscle is equally intimately tied to nervous centres. The beat of the heart, freely separable from central nervous action, proceeds unimpaired after the heart's removal from the body. Similarly the intestinal and other visceral muscles, *e.g.* bladder, continue their rhythmic contractions after rupture of all extrinsic nervous ties. Almost as full is the muscular independence of the blood vessels. Their tonus is in many cases impaired by

severance of their connexions with the neuraxis (central nervous system), yet the impairment is but transient. Thus the contractile activity of these visceral and vascular muscles is fundamentally independent of central nervous influence. Yet, to all these visceral and vascular muscles, including the heart itself, nerves are distributed from the central nervous system, and through these nerves that system does on occasion influence visceral and vascular activity. The nature of the influence thus exerted is in all these cases twofold. Two influences opposite in kind and direction can be exerted, and sometimes one is employed, sometimes the other. The one augments the contractile activity of the muscle, the other diminishes it: the former is termed excitatory, the latter inhibitory; and the nerve fibres which unfold these two influences respectively are distinct from one another.

"In striking contrast with this large measure of independence from the central nervous system shown by the visceral and vascular muscles stands the complete dependence on that system of the skeletal muscles. The contraction of a skeletal muscle is always the expression of a mandate sent to it from the neuraxis. Even with the peripheral nerve-muscle preparation, a stimulus to the nerve throws the muscle into contraction more readily than a stimulus applied directly to itself. After severance of its nerve, a skeletal muscle never, except under artificial stimulation, contracts again. Hence, when removed from the body, a skeletal muscle, unlike the heart, intestine, stomach, bladder, etc., never of itself contracts again. Mere section of its nerve, even when the muscle remains *in situ*, causes it to lapse into a paralytic quietude so profound that its very structure becomes in a short time hardly recognisable as muscle. Its contractive function and its very nutrition are indissolubly dependent on the neuraxial centre which innervates it.

"That stimulation of the nerve of a skeletal muscle readily and regularly elicits contraction of the muscle was known early. The discovery later of nerves augmenting and accelerating the contractions of the visceral and circulatory muscles fell into line with that previous knowledge, and seemed natural enough. But the discovery by the Webers in 1846 that a nerve passing to the heart can stop or inhibit that muscle's contraction seemed so surprising that at first it was by many not accepted. When assured as a fact, it was, however, recognized as a phenomenon of high significance, unlikely to stand alone. Search for inhibitory nerves to other muscles was begun. The muscle of the intestine wall was shown to have an inhibitory nerve (n. splanchnicus, Pflüger, 1857). The ring-musculature of the submaxillary artery and its branches was shown to receive an inhibitory nerve (chorda tympani, Bernard, 1858). Meanwhile, there was always expectation of discovery of inhibitory nerves to skeletal muscle. In 1885, Pavlov discovered inhibitory nerves for the adductor muscles of the bivalve mollusc Anodon; and similarly Biederman, in 1886, inhibitory nerves for the claw muscles of the arthropod Astacus. For the skeletal muscles of vertebrates, however, no such inhibitory nerves have come to light.

"Yet it were strange did the abundant machinery of the vertebrate nervous system provide no means for checking or curbing the contraction of the muscles of its great skeletal muscular congeries. That congeries, with its manifold individual pieces, some diametrically opposed to others, is so complex, and the confusion and waste of effort consequent were its opponent parts to obstruct each other's action would seem so foreign to Nature's usual

harmonious economy, that inhibitory as well as excitatory control appears *a priori* almost a necessity. The negative result of the search for skeletal inhibitories did not quell the expectation that they would ultimately be found; the search was often renewed. It was, however, always a search for direct efferent inhibitory nerves passing to the muscle from the central nervous system on the plan already discovered for visceral and circulatory muscles. Yet when the broad facts are kept in view, that plan does not appear as necessarily the one most likely to exist. With the skeletal muscle, contraction has become so wholly dependent on the central nervous system that it contracts only when the immediate spinal or cranial motor centre summons it to do so. With a muscle which contracts solely at behest of its motor nerve-centre, a simple plan to control the muscle is to control the centre. And, in fact, it is now found that the inhibitory control over skeletal muscle is obtained by centripetal inhibitory nerves acting on the muscles' motor centres; the control is not by centrifugal nerves directly inhibiting the muscles themselves. Hence, unlike the autonomous muscles, visceral and circulatory, which receive two kinds of centrifugal nerves, excitatory and inhibitory, the skeletal muscle receives but one nerve, namely, an excitatory or motor nerve. No inhibitory nerve is supplied to it. But its motor centre receives two kinds of centripetal nerves, excitatory and inhibitory.

"The state of the skeletal muscle reflects faithfully the state of its motor centre. Its motor centre is the only source whence impulses can reach it. Upon that motor centre many nerve-paths converge, transmitting to it nervous impulses from various receptive points and from centres elsewhere. Of these nerve-paths some excite, others inhibit. The latter, by quelling or moderating the discharge from the motor centre, quell or moderate the contraction of the muscle. The inhibition of the skeletal muscle is therefore always reflex; and the study of skeletal inhibition falls wholly under the head of reflex action. The motor centre is a convergence point for various reflex influences competing, so to say, for dominance over the muscle. The motor centre lies as an instrument passive in the hands of opposing forces of excitation and inhibition, exerted by nerve channels which impinge upon it. Sometimes the one influence and sometimes the other influence prevails; often the two are simultaneously in action, and then these opponents partially cancel, either spatially in proportion to the relative intensity of their respective stimulations, or temporally by alternating with each other rhythmically in their dominance over the motor centre, thus producing rhythmic reflexes." [This long quotation is from Sherrington's paper entitled "Reflex Inhibition as a Factor in the Co-ordination of Movements and Postures" in *Quart. J. exp. Physiol.*, 1913, *6*, 251–310.]

INDEX[1]

[1] Prepared by I. Calma and Sheila Wright.

36*